

FIGURE 101

GGGCCGACCATTATCCAACCGGGNTCACTGTTGGCTCATCTCCCTCCTGGATGAANCAGCGC
CATCNTCAGACTCCCTGCCCATGGAGATTTNNCCTATGCTGGCGACAACATCNTGACCCCC
AGCCATGTACGAGGGCTTGAACGTCNGCGTGCAGANCACCAGATCCAGTGCAA
AGTCTTGACTCCTGCTGAATCTGNGCAGCACATTGCAGCAACCCNTGCCCTGATGGTGGT
TGGCATCCTCCTGGAGTGATAAGCAATCTTGTGCCACCGTTGGCATGAAGTGTATGAAGT
GCTTGGAAAGACGATGAGGTGCAGAAGATGAGGATGGCTGTCATTGGGGCGCGATATTCTT
CTTGCAGGTCTGGCTATTNNNNTGCCACAGCATGGTATGGCAATAGAATCGTCAAGAAT
TCTATGACCCTATGACCCCCAGTCAATGCCAGGTACGAATTGGTCAGGCTCTTCAC TGCG
TGGGCTGCTGCTCTCTGCCTCTGGGAGGTGCCCTACTTGCTGTTCTGCGA

702 1237

FIGURE 102

ATTCTCCCCTCCTGGATGGATGGATGCNCCACCGTCACATTGCCTTCCCCANTGGAGGATTNAC
TCCTATGCTGGCGACAACATCGTACCCCCCAGGCCATTACCGAGGGGCTTGGATGTCNT
GCNTGTCGAGAGCACCGGGCAGATCCCAGTGCAAAGTCTTGACTCCTGCTGAATCTGAG
CAGCACATTGCAAGCAACCGTGCCTGATGGGTTGCATCCTCCTGGAGTGATAGCAAC
CTTGTCGCCACCGTTGGCATGAAGTGTATGAAGTGCTTGGAAAGACGATGAGGTGCCAGAAG
ATGAGGATGGCTGTCATTGGGGCGCGATATTCTGTTGCAGGTCTGGCTATTAGTNGC
CACAGCATGGTATGGCAATAGANTNNTCNNNNNTCTATGACCTATGACCCAGTCAATG
CCAGGTACGAATTGGTCAGGCTCTCTTCACTGGCTGGCTGCTGCTTCTCTGCCTCTG
GGAGGTGCCCTACTTGCTGTTCTGTCCC

103 / 237

FIGURE 103

AGAGCACCGGCAGATCCCAGTNCAAAGTCTTGACCCCTGCTGAATCTGAGCAGCACATTNC
AAGCAACCCCTTGCCTGAAGGTGGTTGNCATCCCCCTGGGAGTGAATAGCAATCTTGTG
GCCACCCTGGCATGAAGTNATGAAGTGCTTGGAAAGACGATGAGGTGCAGAAGATGAGGAT
GGCTGTCATTGGGGCGCGATATTCTTGCAGGTCTGGCTATTTAGTNCCACAGCAT
GGTATGGCAATAGNATNNTCGNGGNTCTATGACCCATGACCCAGTCAATGCCAGGTAC
GAATTGGTCAGGCTCTTCACTGGCTGGCTGCTCTCTGCCTCTGGGAGGTGC
CCTACTTGCTGTTCTGTCCCCGAA

1041237

FIGURE 104

AGCAATGCCCTGCCCCCAGTGGAGGATTAATTCTATGNTGGGGACAACATTGTGACNGCC
AGGCCATGTACGGGGGCTGTGGATGTCCTGCCTGCAGAGCACCGGGCAGATCCAGTGC
AAAGTNTTGACTCCTGCTGAATTGAGCAGCACATTGCAAGCAACCCGTGCCTGATGGT
GGTTGGCATCTCCTGGGAGTGATAGCAATCTTGTGCCACCGTGGNAATGAAGTGTATGA
AGTGCTTGAAGACGATGAGGTGCAGAAGATGAGGATGGCTGTCAATTGGGGCGCAGATATT
CTTNTTGCAAGGTCTGGCTATTTAGTTGCCACAGCATGGTATGGCAATAGAATNGTTCAAGA
ATTTTATGACCCTATGACCCCAGTCAATGCCAGGTACGAATTGGTCAGGCTTNTTCAC TG
GCTGGGCTGCTGCTNTTCTGCCNTGGGAGGTGCCCTANTTGCTGTTCTGCGAAC C

1051237

FIGURE 105

TCATAGGGGGCGCGATATTTTCTTGCAGGTNTGGTTATTTAGTTGCCACAGCATGGTA
TGGCAATAGAACGTTCAAGAATTNTATGACCCATGACCCCAGTCAATGCCAGGTACGAAT
TTGGTCAGGCCTCTNTTCACTGGNTGGCTGCTGCTCTNTNGCCTNTGGGAGGTGCCCTA
CTTGCTGTTCTG

106/237

FIGURE 106

TTCCTGGGATGGATCCGCCCATCNTCACATGCCCTGCCCNNTGGAGATTACNCCTATGC
TGGCGAACAAACATCNTGACCGCCCAGGCCATGTACGAGGGCTGTGGAATGTCCCTGCGTGTG
CCAGAGCACCGGGCAGATCCAGTGCAAAGTCTTGACTCCTGCTGAATCTGAGCAGCACAT
TGCAAGCAACCNTGCCTGATGGTGGTTGGCATCCTCCTGGGAGTGATAGCAATCTTGTGG
CCACCGTTGGCATGAAAGTGTATGAAGTGCTTGGAAAGACGATGAGGTGCAGAAGATGAGGAT
GGCTGTCATTGGGGCGCGATATTCTTCTTGCAAGGTCTGGCTATTTAGNNCCACAGCAT
GGTATGGCAATCAGACCCNNTCANAAACTCTATGACCCATGACCCAGTCAATGCCAGGTA
CGAATTGGTCAGGCCTCTTCACTGGCTGGCTGCTGCTCTGCCTCTGGGAGGTG
CCCTACTTGCTGTTCCGTCCCCGAAAAACAACCTTTACCCACG

107/237

FIGURE 107

CGGGGCTGCAGCTGTTGGCTTCATCTCGCTTCCCTGGATGGAATCGCGGCCATCGTCAGCA
CTGCCCTGCCCATGGAGGATTACTCNTATGCTGGCGACAACATCGTGACCNCCCAGGCCA
TGTACGAGGGCTGTGGATGTCNGCGTGTGCAGAGCACCGGGCAGATCCAGTGCAGTCT
TTGACTCCTGCTGAATCTGAGCAGCACATTGCAAGCAACCNTGCCTTGTGGTGGTGGCA
TCCTCCTGGAGTGATAGCAATCTTGTGGCCACCCTGGCATGAAGTGTATGAAGTGCTTG
GAAGACGATGAGGTGCAGAAGATGAGGATGGCTGTCAATTGGGGCGCGATATTCTTCTTGC
AGGTCTGGCTATTNTAGTTGCCACAGCATGGTATGGCAATAGAATCGTTCAAGAATTCTAT
GACCCTATGACCCAGTCAATGCCAGGTACGAATTGGTCAGGCTCTTCACTGGCTGGC
TGCTGCTTCTCTGCCTCTGGAGGTGCCCTACTTGCTGTTCTGCGAA

108/237

FIGURE 108

GC GTGCCGT CAGCTCGCCGGG CACCGCGGCCCTCGCCCTCGCCCTCCGCCCTGCGCCTGCAC
CGCGTAGACCGACCCCCCCCCTCCAGCGCCCACCCGGTAGAGGACCCCCGCCGTGCCCG
ACCGGTCCCCGCCCTTTGTAAAACCTAAAGCGGGCGCAGCATTAAACGCTTCCGCCCGT
GACCTCTCAGGGGTCTCCCCGCAAAGGTGCTCCGCCGCTAAGGAACATGGCGAAGGTGGAG
CAGGTCTGAGCCTCGAGCGCAGCACGAGCTCAAATTCCGAGGTCCCTCACCGATGTTGT
CACCAACCTAAAGCTTGCAACCCGACAGACCGAAATGTGTGTTAAGGTGAAGACTA
CAGCACCACGTAGGTACTGTGTGAGGCCAACAGCGGAATCATCGATGCAGGGCCTCAATT
AATGTATCTGTGATGTACAGCCTTCGATTATGATCCAATGAGAAAAGTAAACACAAGTT
TATGGTCAGTCTATGTTGCTCCAATGACACTTCAGATATGGAAGCAGTATGGAAGGAGG
CAAACCGGAAGACCTTATGGATTCAAACCTTAGATGTGTGTTGAATTGCCAGCAGAGAAT
GATAAACACATGATGTAGAAATAAATAAATTATATCCACAATGCAAAAGACAGAAAC
ACCAATAGTGTCTAAGTCTGAGTTCTTGGATGACACCGAAGTTAAGAAGGTTATGG
AAGAATGTAAGAGGCTGCAAGGTGAAGTTCAGAGGCTACGGGAGGAGAACAGCAGTTCAAG
GAAGAAGATGGACTGCGGATGAGGAAGACAGTGCAGAGCAACAGCCCCATTTCAGCATTAGC
CCCAACTGGGAAGGAAGAACGGCTTAGCACCCGGCTTGGCTCTGGGTTGTTCTTA
TCGTTGGTGTAAATTGGGAAGATTGCCCTG**TAGAGGTAGCATGCACAGGATGGTAAATTG**
GATTGGTGGATCCACCATATCATGGGATTAAATTATCATAACCATGTGTAAAAAGAAATT
AATGTATGATGACATCTCACAGGTCTGCCTTAATTACCCCTCCCTGCACACACATACAC
AGATACACACACACAAATATAATGTAACGATCTTTAGAAAGTTAAAATGTATAGTAACG
ATTGAGGGGGAAAAAGAATGATCTTATTAAATGACAAGGGAAACCATGAGTAATGCCACAAT
GGCATATTGTAATGTCATTAAACATTGGTAGGCCCTGGTACATGATGCTGGATTACCTC
TCTTAAATGACACCCTCCTCGCCTGTTGGTCTGGCCCTGGGGAGCTGGAGCCCAGCAT
GCTGGGGAGTGCAGCTCACAGTAGTCCCACGTGGCCACTCCGGCCAGGCTG
CTTCCGTGTTCAAGCCATCAGCTCTGGACTGATGAACAGAGTCAGA
AGCCCAAAGGAATTGCACTGTGGCAGCATCAGACGTACTCGTCATAAGTGAGAGGCGTGT
TGACTGATTGACCCAGCGCTTGGAAATAATGGCAGTGTGTTGTTCACTAAAGGGACCAA
GCTAAATTGATTGGTCATGTAGTGAAGTCAAACTGTTATTCAAGAGATGTTAATGCATA
TTAACTTATTAAATGTATTTCATCTCATGTTCTTATTGTCACAAGAGTACAGTTAATGC
TGCCTGCTGCTGAACTCTGTTGGGTGAACTGGTATTGCTGCTGGAGGGCTGGGCTCCTCT
GTCTCTGGAGAGTCTGGTCATGTGGAGGTGGGTTATTGGGATGCTGGAGAAGAGCTGCCA
GGAAGTGTGTTCTGGTCAGTAAATAACAACTGTCATAGGGAGGGAAATTCTCAGTAGTG
ACAGTCAACTCTAGGTTACCTTTAAATGAAGAGTAGTCAGTCTTAGATTGTTCTTATA
CCACCTCTCAACCATTACTCACACTCCAGCGCCCAGGTCCAAGTCTGAGCCTGACCTCCCC
TTGGGGACCTAGCCTGGAGTCAGGACAAATGGATCGGGCTGCAGAGGGTTAGAAGCGAGGGC
ACCAGCAGTTGAGAGTTGACTGTGAATTAAATTATGCCATAAAAGACCAACCCAGTTCTGTTGA
CTATGTAGCATCTGAAAAGAAAAATTATAATAAGCCCCAAAATTAAAGAAAA

109/237

FIGURE 109

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53977
<subunit 1 of 1, 243 aa, 1 stop
<MW: 27228, pI: 7.43, NX(S/T): 2
MAKVEQVLSLEPQHELKFRGPFTDVVTNLKGNTDRNVCFKVKTAPRRYCRPNNSGIID
AGASINVSMQLQPFDYDPNEKSCHKFMQSMFAPTDTSDMEAVWKEAKPEDLMDSKLRCVFE
LPAENDKPHDVEINKIISTTASKTETPIVSKSLSSLDTEVKVMECKRLQGEVQRLREE
NKQFKEEDGLRMRKTVQNSPISALAPTGKEEGLSTRLLALVVLFIVGVIIIGKIAL

Important features:**Transmembrane domain:**

amino acids 224-239

N-glycosylation site.

amino acids 68-71

N-myristoylation site.

amino acids 59-64, 64-69 and 235-240

1101 237

FIGURE 110

GTCAGTCTCTAGATTGTCCTTATCCCACCTTCAACCANTACTCACATTCNAGGCCAG
GTCCANGTCTGAGCCTGACTTCCCCTGGGGACCTAGCCTGGAGTCAGGACAATGGNTCGGG
CTGCAGAGGNTTAGAAGCGAGGGCACCAGCAGTTGGGTGGGGAGCAAGGGNNAGAGAAA
CTCTCAGCGAACCTCTAGTACTAGTTGAGAGTTGACTGTGAATTAATTTATGCCATA
AAAGACNAACCCAGTTCTGTTGACTATGTAGCATCTGAAAAGAAAAATTATAATAAGCC
CCAAAATTAAGAATTCTTGTCACTTGCTATGGGGGAATTATTATTTT
ATCATTTTATTATTTGCCATTGGAAGGTTAACTTAAAATGAGC

1111237

FIGURE 111

TATTGTAAAGGCCATTTAAACCATTGGTAGGCCTTGGTACATGATGCTGGATTACCTCCTT
AAATGACACCNTTCCTCGCCTGTTGGTCTGGCCNTTGGGGAGCTGGAGCCCCAGCATGCTG
GGGAGTGCGGTCACTCCACACAGTAGTCCCCACGTGGCCACTCCGGCCCAGGCTGCTTT
CCGTGTCTTCAGTTCTGTCCAAGCCATCAGCTCCTGGACTGATGAACAGAGTCAGAACGCC
CAAAGGAATTGCCACTGTGGCAGCATTACAGTACTCGTATAAGTGAGAGGCGTGTGTTGA
CTGATTGACCCAGCGCTTGGAAATAATGGCAGTGCTTGTTCACTAAAGGACCAAGCT
AAATTGTATTGGTCATGTAGTGAAGTCAAACGTATTCAAGAGATGTTAATGCATATTAA
ACTTATTAAATGTATTCACTCATGTTTCTTATTGTCACAAGAGTACAGTTAATGCTGCG
TGCTGCTGAACCTGTTGGGTGAACGGTATTGCTGGAGGGCTG

7121237

FIGURE 112

CCCTGGTGGTTTGTCTTAATCGTTGGTGAATTNTGGGAAGATTGCTTAGAGGTA
GNATGCACCNGGCTGGTAAATTGGATTGGTGGATCCACCATATCCATGGGATTAAATTAT
CATAACCATGTGTAAAAAGAAATTAATGTATGATGACATNTCACAGGTATTGCCTTAAATT
ACCCATCCCTGNANACACATACACAGATACACANANACAAATNTAATGTAACGATNTTTAG
AAAGTTAAAATGTATAGTAAC

713/237

FIGURE 113

GGTGGCCCATTCCCGGCCAGGCTGCTTCGGTNTTCAGTTCTGTCCAAGCCATCAGCTCC
TTGGGACTGATGAACAGAGTCAGAAGCCAAAGGAATTGCACTGTGGCAGCATNAGACGTAC
TTGTNATAAGTGAGAGGCCTGTTGACTGATTGACCCAGCGCTTGAAATAATGGCAGT
GCTTGTTCACTAAAGGGACCAAGCTAAATTGATTGGTCATGTAGTGAAGTCAAACTG
TTATTCAGAGATGTTAACATGCATATTAANTTAAATGTATTTNATNTCATGTTTCTTA
TGTCACAAGAGTACAGTTAACATGCTGCGTGCTGCTGAANTNTGTTGGGTGAACGGTATTGC
TGCTGGAGGGCTGTGGCTCCTCTGTCTTGGAGAGTCTGGTCATGTGGAGGTGGG

114/237

FIGURE 114

TGCTTCGGTCTTCAGTTCTGTCCAAGCCATCAGCTCCTGGGACTTGATGAACAGAGTC
AGAACGCCAAAGGAATTGCACTGTGGCAGCATCAGACGTACTCGTCATAAGTGAGAGGCCTG
TGTTGACTGATTGACCCAGCGCTTGGAAATAAATGGCAGTGCTTGTTCACTAAAGGGAC
CAAGCTAAATTGTATTGGTCATGTAGTGAAGTCAAACGTGTTATTCAAGAGATGTTAATGC
ATATTTAACTTATTAATGTATTCATCTCATGTTCTTATTGTCACAAGAGTACAGTTAA
TGCTGCGTGC

1151237

FIGURE 115

AAACCTTAAAAGTTGAGGGAAAAGAATGATCCTTATTAAATGACAAGGGAAACCNTGN
AATGCCACAATGGCATATTGTAAATGTCATTTAACATTGGTAGGCCTTGGTACATGATGC
TGGATTACCTCTCTTAAAATGACACCCCTCCTCGCCTGTTGGTGCTGGCCCTGGGAGCTN
GAGCCCAGCAGTGGGGAGTGCAGTCTGCTCCACACAGTAGTCCCCANGTGGCCANTCCC
GGCCCAGGCTGCTTCCGTCTTCAGTTCTGTCAGCCATCAGCTCCTGGGANTGATGA
ACAGAGTCAGAAGCCAAAGGAATTGCANTGTGGCAGCAGTCAAGTCAAGTGA
GAGGCAGTGTGTTGANTGATTGACCCAGCGCTTGGAAATAATGGCAGTGCCTTGTCANTT
AAAGGGNCCAAGNTAAATTGTATTGGTCATGTAGTGAAGTCAAANTGTTATTCAGAGATG
TTAATGCATATTAANTTAAATGTATTCATNTCATGTTTCTTATTGTACAAGGGT
ACAGTTAATGCTGCGTGTGAANTCTGTTGGTGAANTGGTATTGCTG

7161237

FIGURE 116

GGCCCTGGGAGCTGGAGCCCAGCATGCTGGGAGTCGGTCAGCTCCACACAGTAGTCCC
CACGTGGCCCACCTCCGGCCAGGCTGCTTCCGTGTCTCAGTTCTGTCAGGCCATCAGC
TCCTTGGACTGATGAACAGAGTCAGAACGCCAAAGGAATTGCACTGTGGCAGCATCAGACG
TACTCGTCATAAGTGAGAGGCCTGTTGACTGATTGACCCAGCGCTTGGAAATAAATGGC
AGTGCTTGTTCACTTAAAGGGACCAAGCTAAATTGTATTGGTCATGTAGTGAAGTCAAA
CTGTTATTCAAGAGATGTTAATGCATATTAACCTATTAAATGTATTCATCTCATGTTTC
TTATTGTCACAAGAGTACAGTTAATGCTGCGTGCTGCTGAACCTGTTGGGTGAACGGTAT
TGCTGCTGGAGGGCTGTGGCTCCTCTGCTCTGGAGAGTCGGTCATGTGGAGGTGG

777237

FIGURE 117

GCGAGCTCCGGGTGCTGTGGCCCCCTGGGGGGCCTCCGGCTCAGGCTGGCTGAGA
 GGCTCCCAGCTGCAGCGTCCCCGCCCTCCTCGGGAGCTCTGATCTCAGCTGACAGTGCC
 CTCGGGGACAAACAAGCCTGGCAGGGCTCACTTGTGCCCAGGCTGGAGTTAGTGCC
 TGATCATGGTTACTGCAGCCTGACCTCCTGGGTTCAAGCGATCCTGCTGAGTAGCTGGGA
 CTACAGGACAAAATTAGAAGATCAAATGGAAGATGGGATTTATGTGGCACTTGAGAAA
 CCTGGGTGGACCCCTATTGATGGATCTGAAATGGAATGGGATTTATGTGGCACTTGAGAAA
 GGTACCCCCGATTGTCAGTGAAAGGACTTCCATCTCACCAAGCCCCGATTGAGGCAGATG
 CTAAGATGATGGTAAATACAGTGTGGCAGTGAATGCCAGAAAGAACTCCAACTCCAGC
 CTTCTGAATTGGAGGATTATCTTCCTATGAGACTGTCTTGAGAATGGCACCCGAACCTT
 AACCAAGGGTGAAAGTTCAAGATTGGTCTTGAGCCGACTAAAATATCACCAACAAAGGGAG
 TATCTGTTAGGAGAAAGAGACAGGTGTATGGCACCGACAGCAGGTTAGCAGTCTTGAGAAA
 AGGTTCTTAACCAATTCCCTTCAGCACAGCTGTGAAGCTTCCACGGGCTGTAGTGGCAT
 TCTCATTCCCTCAGCATGTTCTAAGTGTCTGCCCAGTGTGTTCATGATGGAAAGGACTATG
 TCAAAGGGAGTAAAAGCTAAGGGTAGGGTTGTGAAGATGAGGAATAAAAGTGGAGGCAAG
 AACAGTCGAGGTTCTAACAGAGGAGCAGGAGAGAAGCTAGTGGTGGTGAACAAAGAGAGGGTAC
 CAGAGAGCATCTGCAGGAGAGAGCGAAGGGTGGGAGAAGAAGAAAAAAATCTGCCGGGTC
 AGAGGATTGCCGAAGGGAGGCCTCCTTCAGTGGACCCGGGTCAGAATACCCACATTCCG
 AAGGGCTGGCACGAGGAGGCATGGGGACGCTACCTGGACTATGACTATGCTCTTGGA
 GCTGAAGCGTGCACAAAAAGAAATACATGGAACCTGGAAATCAGCCAACGATCAAGAAAA
 TGCCTGGTGAATGATCCACTTCAGGATTTGATAACGATAGGGCTGATCAGTTGGTCTAT
 CGGTTTGCAGTGTGCGACGAATCCAATGATCTCCTTACCAATACTGCGATGCTGAGTC
 GGGCTCCACCGGTCGGGGCTATCTGCGTCTGAAAGATCCAGACAAAAAGAATTGGAAGC
 GAAAATCATTGCGGCTACTCAGGGCACCGAGTGGTGGATGTCACGGGTTAGAAGGAC
 TACAACGTTGCTTCGCATCAGTCCCTAAATACGCCAGATTGCTCTGGATTACCG
 GAACGATGCCAATTGCGCTACGGCTAACAGAGACCTGAAACAGGGCGGTGTATCATCTAA
 TCACAGAGAAAACCAGCTCGCTTACCGTAGTGAGATCACTCATAGTTATGCTGGACTT
 GAACTCTGTCATAGCATTCAACATTTCAGGAGATTTCGTCATTAACTCTTCA
 TGTATAGGTGCAGATATTGAAACTAGGTGGGCACTTCAGTCAATGCCAAGTATATACTCTT
 CATGGTGATGAGTTCTATTGAGAAAATTGCTCTTAAATAGACACACTT
 AACACCTCAACAGGTATTATAAAACATGTGACTCCTTAATGGACTTATTCTCAGGGTCC
 TACTCTAACAGAATCTAACAGGATGCTGGTGTGTTAAATGTGAAATTGCATAGATAAA
 GGTAGATGGTAAAGCAATTAGTATCAGAATAGAGACAGAAAGTTACAACACAGTTGACTA
 CTCTGAGATGGATCCATTGCTCATGCCCTCAATGTTATATTGTTATCTGTTGGGTCT
 GGGACATTAGTTAGTTAGGAAAGATTACAAATCAGAAGAAAAGCAAGCATTATAAA
 CAAAACAAACTCTTAACTGCTTAAAGAAATAACAATTACAATGTGTTAAATTTAAAAAA
 TGGGAGAAATAGTTGTTCTATGAAATAACCTAGTTAGAAATAGGGAGCTGAGACATT
 TAAGATCTCAAGTTTATTAACTAATACTCAAAATATGGACTTTCATGTATGCATAGGG
 AAGACACTTCACAAATATGAATGATCATGTGTTGAAGGCCACATTATTATGCTATACAT
 TCTATGTATGAGGTGCTACATTAGGACAAAGAATTCTGTAATCTTTCAAGAAAGAGT
 CTTTTCTCCTTGACAAAATCCAGCTTGTATGAGGACTATAGGGTGAATTCTCTGATTAG
 TAATTAGATATGTCCTTCTAAATGAATAAAATTATGAATATGA

718/237

FIGURE 118

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57253
<subunit 1 of 1, 413 aa, 1 stop
<MW: 47070, pI: 9.92, NX(S/T): 3
MENMLLWLIFFTPGWTLIDGSEMEWDFMWHLRKVPRIVSERTFHLTSPAFEADAKMMVNVC
GIECQKELPTPSLSELEDYLSYETVFENGTRTLTRVKVQDLVLEPTQNITTGVSVRRKRQV
YGTDSRFSILDKRFLTNFPFSTAVKLSTGCSGILISPQHVLTAAHCVHDGKDVKGSKKLRV
GLLKMRNKGKRRGSKRSRREASGGDQREGTREHLQERAKGGRRKKSGRGQRIAEGRPS
FQWTRVKNTHIPKGWARGGMGATLDYDYALLELKRAHKKKYMELGISPTIKKMPGGMIHFS
GFDNDRADQLVYRFCVSDESNDLLYQYCDAESGSTGSGVYRLKDPKKNWKRKIIAVYSG
HQWVDVHGVQKDYNNAVRITEPLKYAQICLWIHGNDANCAYG
```

Important features:**Signal peptide:**

amino acids 1-16

N-glycosylation sites.

amino acids 90-93, 110-113 and 193-196

Glycosaminoglycan attachment site.

amino acids 236-239

Serine proteases, trypsin family, histidine active site.

amino acids 165-170

7191237

FIGURE 119

AATGTGAGAGGGGCTGATGGAAGCTGATAGGCAGGACTGGAGTGTAGCACCAGTACTGGAT
GTGACAGCAGGCAGAGGAGCACTTAGCAGCTTATTCACTGTCCGATTCTGATTCCGGCAAGG
ATCCAAGCATGGAATGCTGCCGTGGCAACTCCTGGCACACTGCTCCTCTTCTGGCTTC
CTGCTCCTGAGTTCCAGGACCGCACGCTCCGAGGAGGACGGGACGGCTATGGGATGCCTG
GGGCCATGGAGTGAATGCTCACGCACCTGCAGGGGGAGGGCCTACTCTGAGGCCT
GCCTGAGCAGCAAGAGCTGTGAAGGAAGAAATATCCGATAACAGAACATGCAGTAATGTGGAC
TGCCCACCAGAAGCAGGTGATTCCGAGCTCAGCAATGCTCAGCTCATAATGATGTCAAGCA
CCATGGCCAGTTTATGAATGGCTTCGTCTAATGACCCCTGACAACCCATGTTCACTCA
AGTGCCAAGCAAAGGAACAAACCTGGTTGTAAGTACAGCACCTAACGGCTTAGATGGTACG
CGTTGCTATACAGAATCTTGGATATGTGCATCAGTGGTTATGCCAAATTGTTGGCTGCGA
TCACCAAGCTGGGAAGCACCGTCAAGGAAGATAACTGTGGGCTGCAACGGAGATGGTCCA
CCTGCCGGCTGGTCCGAGGGCAGTATAAATCCCAGCTCCGCAACCAAATGGATGATACT
GTGGTTGCACCCCTATGGAAGTAGACATATTGCCTGTCTAAAAGGTCTGATCACTT
ATATCTGAAACCAAACCCCTCCAGGGACTAAAGGTGAAAACAGTCTCAGCTCACAGGAA
CTTCCTGTGGACAATTCTAGTGTGGACTCCAGAAATTCCAGACAAAGAGATACTGAGA
ATGGCTGGACCACTCACAGCAGATTTCATTGTCAAGATTGTAACCTGGGCTCCGCTGACAG
TACAGTCCAGTTCATCTTCTATCAACCCATCATCCACCGATGGAGGGAGACGGATTCTTC
CTTGCTCAGCAACCTGTGGAGGAGGTATCAGCTGACATCGGCTGAGTGCTACGATCTGAGG
AGCAACCGTGTGGTGTGACCAAAACTGTCACTATTACCCAGAGAACATCAAACCCAAACC
CAAGCTTCAGGAGTGCAACTTGGATCCTGTCCAGCCAGTGACGGATAACAGCAGATCATGC
CTTATGACCTCTACCATCCCCTCCTCGGTGGAGGCCACCCATGGACCGCGTGCTCCTCC
TCGTGTGGGGGGGCATCCAGAGGCCAGTTCTGTGTGGAGGAGGACATCCAGGGCA
TGTCACTTCAGTGGAAAGAGTGGAAATGCATGTACACCCCTAACATGCCATCGCGCAGCCCT
GCAACATTTTGACTGCCCTAAATGGCTGGCACAGGAGTGGCTCCGTGCACAGTGACATGT
GGCCAGGGCCTCAGATACCGTGTGGCCTCTGCATCGACCATCGAGGAATGCACACAGGAGG
CTGTAGCCCCAAAACAAAGCCCCACATAAAAGAGGAATGCATCGTACCCACTCCCTGCTATA
AACCCAAAGAGAAAATTCAGTCAGGCCAAGTTGCCATGGTTCAAACAGCTCAAGAGCTA
GAAGAAGGAGCTGCTGTCAAGAGGAGCCCTCGTAAGTGAAAAGCACAGACTGTTCTATA
TTTGAAAACTGTTTGTAAAGAAAGCAGTGTCTCACTGGTTGTAGCTTCATGGTTCTGA
ACTAAGTGTAAATCATCTCACCAAGCTTTGGCTCTCAAATTAAAGATTGATTAGTTCAA
AAAAAAA

720/237

FIGURE 120

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA58847
<subunit 1 of 1, 525 aa, 1 stop
<MW: 58416, pI: 6.62, NX(S/T): 1
MECCRATPGTLLLFLAFLLLSSRTARSEEDRDGLWDAWGPWSECSRTCGGGASYSLRRCLS
SKSCEGRNIRYRTCSNVDCPPEAGDFRAQQCSAHNDVKHHGQFYEWLPVSNDPDNPCSLKCQ
AKGTTLVVELAPKVLGTRCYTESLDMCISGLCQIVGCDHQ LGSTVKEDNCGVNGDGSTCR
LVRGQYKSQLSATKSDDTVVALPYGSRHIRLVLKGPDHLYLETKTLQGTKGENSLSSTGTFL
VDNSSVDFQKFPDKEILRMAGPLTADFIVKIRNSGSADSTVQFIFYQPIIHRWRETdffPCs
ATCGGGYQLTSAECYDLRSNRVVADQYCHYY PENIKPKPKLQECNLDP PASDGYKQIMPYD
LYHPLPRWEATPWTACSSSCGGGIQSRAVSCVEEDIQGHVT SVEEWKCMYTPKMPIAQPCNI
FDCPKWLAQEWPCTVTCGQGLRYRVVL CIDH RGMHTGGCSPKT KPHI EECIVPTPCYKPK
EKLPVEAKLPWFQAOELEEGAAVSEEPS
```

Important features:**Signal peptide:**

amino acids 1-25

N-glycosylation site.

amino acids 251-254

Thrombospondin 1

amino acids 385-399

von Willebrand factor type C domain proteins

amino acids 385-399, 445-459 and 42-56

7211237

FIGURE 121

CGGACGCGTGGCGGCGCTGCGGAACCTCCGTGGAGGGCCGGTGGCCCTCGGGCTGAC
AGATGGCAGTGGCCACTGCGGCGCAGTACTGGCCCTCTGGGCGGGCGCTGTGGCTGGCG
GCCCGCCGGTTCGTGGGCCAGGGTCCAGCGGCTGCGCAGAGGCGGGACCCGGCTCAT
GCACGGGAAGACTGTGCTGATCACCGGGCGAACAGCAGGCTGGCCGCCACGGCGCCG
AGCTACTGCGCTGGGAGCGCGGGTGATCATGGCTGCCGGACCAGCGCGCGCGAGGAG
GCGGCGGGTCAGCTCCGCCGAGCTCCGCCAGGCCGGAGTGCAGGCCCAGAGCCTGGCGT
CAGCGGGTGGCGAGCTCATAGTCCGGAGCTGGACCTCGCTCGCTCGCTCGGTGCG
CCTCTGCCAGGAATGCTCCAGGAAGAGCCTAGGCTGGATGTCTGATCAATAACGAGGG
ATCTCCAGTGCCCTTACATGAAGACTGAAGATGGGTTGAGATGCAGTCGGAGTGAACCA
TCTGGGCACTTCTACTCACCAATCTCTCCTGGACTCCTCAAAAGTTCAGCTCCCAGCA
GGATTGTGGTAGTTCTCAAACCTTATAAATACGGAGACATCAATTGATGACTTGAAC
AGTGAACAAAGCTATAATAAAAGCTTTGTTAGCCGGAGCAAACGGCTAACATTCTTT
TACCAGGGAACTAGCCCGCCGCTTAGAAGGCACAAATGTCACCGTCAATGTGTTGCATCCTG
GTATTGTACGGACAAATCTGGGGAGGCACATACACATTCCACTGTTGGTCAAACCACTCTC
AATTGGTGTCACTGGCTTTTCAAACCTCCAGTAGAAGGTGCCAGACTTCCATTATT
GGCCTCTCACCTGAGGTAGAAGGAGTGTCAAGGAGATACTTGGGATTGTAAGAGGAAG
AACTGTTGCCAAAGCTATGGATGAATCTGTTCAAGAAAACGACTGGGATATCAGTGAAGTG
ATGGTTGGCCTGCTAAATAGGAACAAGGAGTAAAGAGCTGTTATAAAACTGCATATCAG
TTATATCTGTGATCAGGAATGGTGTGGATTGAGAACTTGTACTTGAAGAAAAAGAATTG
ATATTGGAATAGCCTGCTAAGAGGTACATGTGGGTATTTGGAGTTACTGAAAAATTATT
TGGGATAAGAGAATTCAGCAAAGATGTTAAATATATAGTAAGTATAATGAATAATAA
GTACAATGAAAAATACAATTATATTGTAAAATTATAACTGGGCAAGCATGGATGACATATTA
ATATTGTCAGAATTAAGTGAUTCAAAGTGTATCGAGAGGTTTCAAGTATCTTGAGTT
TCATGCCAAAGTGTAACTAGTTTACTACAATGTTGGTGTGGAAATTATCTGC
CTGGTGTGTGCACACAAGTCTTACTGGAATAAATTACTGGTAC

722 1237

FIGURE 122

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA58747
<subunit 1 of 1, 336 aa, 1 stop
<MW: 36865, pI: 9.15, NX(S/T): 2
MAVATAAAVLAALGGALWLAARRFVGPRVQRLRRGGDPGLMHGKTVLITGANSGLGRATAAE
LLRLGARVIMGCRDRARAEEAAGQLRRELQRQAAECGPEPGVSGVGELIVRELDLASLRVRA
FCQEMLQEEPRLDVLINNAGIFQCPYMKTEDGFEMQFGVNHLGHFLLTNLLLGLKSSAPSR
IVVVSSKLYKYGDINFDDLNSEQSYNKSFCYSRSKLANILFTRELARRLEGTVNVNLHPG
IVRTNLGRHIHIPLLVKPLFNLVSWAFFKTPVEGAQTSIYIASSPEVEGVSGRYFGDCKEEE
LLPKAMDESVARKLWDISEVMVGLLK

Important features:**Signal peptide:**

amino acids 1-21

Short-chain alcohol dehydrogenase family protein

amino acids 134-144, 44-56 and 239-248

N-glycosylation site.

amino acids 212-215 and 239-242

723/237

FIGURE 123

GGGGATTGTAAAGAGGAAGNACTGTGCCAAAGNTATGGATGAATCTGTTGCAAGAAAATTN
TGGGATATCAGTGAAGTGTGGTTNGCCTGCTAAAATAGGAACAAGGAGTAAAAGAGCTGTT
TATAAAACTGCATATCAGTTATCTGTGATCAGGAATGGTGTGGATTGAGAACCTGTTACT
TGAAGAAAAAGAATTTGATATTGGAATAGCCTGNTAAGAGGNACATGTGGGTATTTGGAG
TTACTGAAAAATTATTTGGGATAAGAGAATTCAGCAAAGATGTTAAATATATAGT
AAGTATAATGAATAATAAGTACAATGAAAAATACAATTATATTGTAAAATTATAACTGGCA
AGCATGGATGACATATTAATATTGTCAGAATTAAGTGACTCAAAGTGTATCGAGAGGTT
TTCAAGTATCTTGAGTTCATGCCAAAGTGTAACTAGTTTACTACAATGTTGGTGT
TGTGTGGAAATTATCTGCCTGGCTT

7241237

FIGURE 124

GAGAGGACGAGGTGCCGTCGCTGGAGAACCTCCGCTGCCGTCGGCTCCGGAGCCCAGCC
CTTCCCTAACCAACCCAACCTAGCCCAGTCCCAGCCAGCGCCTGTCCTGTCACGGAC
CCCAGCGTTACC**ATG**CATCCTGCCGTCTCCTATCCTACCCGACCTCAGATGCTCCCTCT
GCTCCTGGTAACTTGGGTTTTACTCCTGTAACAACAGAAATAACAAGTCTTGCTACAGAGA
ATATAGATGAAATTAAACAATGCTGATGTTGCTTAGTAAATTATGCTGACTGGTGT
CGTTTCAGTCAGATGTTGCATCCAATTTGAGGAAGCTCCGATGTCATTAAGGAAGAATT
TCCAAATGAAAATCAAGTAGTGTGTTGCCAGAGTGTGATCAGCACTCTGACATAGCCC
AGAGATAACAGGATAAGCAAATACCAACCCCTCAAATTGTTGTAATGGGATGATGATGAAG
AGAGAAATACAGGGGTCA CGATCAGTGAAGCATTGGCAGATTACATCAGGCAACAAAAAG
TGACCCCATTCAAGAAATTGGGACTTAGCAGAAATCACCCTTGATCGCAGCAAAAGAA
ATATCATTGGATATTGAGCAAAAGGACTCGGACAACACTATAGAGTTTGACCGAGTAGCG
AATATTTGCATGATGACTGTGCCTTCTTCTGCATTGGGATGTTCAAAACCGGAAAG
ATATAGTGGCAGAACATAATCTACAAACACCAGGGCATTCTGCTCCGGATATGGTGTACT
TGGGAGCTATGACAAATTGATGACTTACAATTGGATTCAAGATAATGTGTTCCCTCTT
GTCCGAGAAATAACATTGAAAATGGAGAGGAATTGACAGAAGAAGGACTGCCTTTCTCAT
ACTCTTCACATGAAAGAAGATAACAGAAAGTTAGAAATATTCCAGAATGAAGTAGCTCGGC
AATTAATAAGTAAAAAGGTACAATAAACCTTTACATGCCATTGTGACAAATTAGACAT
CCTCTTCTGCACATACAGAAAATCCAGCAGATTGTCCTGTAATCGCTATTGACAGCTTCT
GCATATGTATGTGTTGGAGACTTCAAAGATGTATTAATTCCGGAAAATCAAGCAATTG
TATTTGACTTACATTCTGGAAAATGACAGAGAATTCCATCATGGACCTGACCCAACTGAT
ACAGCCCCAGGAGAGCAAGCCAAAGATGTAGCAAGCAGTCCACCTGAGAGCTCCTCCAGAA
ACTAGCACCCAGTGAATATAGGTATACTCTATTGAGGGATCGAGATGAGCTT**TAAAAACTTG**
AAAAACAGTTGTAAGCCTTCAACAGCAGCATCACCTACGTGGTGGAAATAGTAAACCTA
TATTTCTATAATTCTATGTGATTTTATTTGAATAAACAGAAAGAAATTAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

725/237

FIGURE 125

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57689
<subunit 1 of 1, 406 aa, 1 stop
<MW: 46927, pI: 5.21, NX(S/T): 0
MHPAVFLSLPDLCSSLVWVFTPVTEITSLENIDEILNNADVALVNFYADWCRFSQ
MLHPIFEEASDVIKEEFPNENQVVFARVDCDQHSDIAQRYSKPTLKLFRNGMMMKREYR
GQRSVKALADYIRQQKSDPIQEIRDLAETTLDRSKRNIIGYFEQKDSDNYRVFERVANILH
DDCAFLSAFGDVSKPERYSGDNIIYKPPGHSAPDMVYLGAMTNFDVTYNWIQDKCVPLVREI
TFENGEEELTEEGLPFLILFHMKEDTESLEIFQNEVARQLISEKGTINFLHADCDKFRHPLLH
IQKTPADCPVIAIDSFRHMYVFGDFKDVLIPGKLKQFVFDLHSGKLHREFHHGPDTDTAPG
EQAQDVASSPPESSFQKLAPSEYRYTLLDRDEL

Important features:**Signal peptide:**

amino acids 1-29

Endoplasmic reticulum targeting sequence.

amino acids 403-406

Tyrosine kinase phosphorylation site.

amino acids 203-211

Thioredoxin family proteins

amino acids 50-66

7261237

FIGURE 126

ATTAAGGAAGAATTCCAAATGAAAATCAAGTAGTNTTGCCAGAGTNGATTGTGATCAGCA
CTCTGACATAGCCCAGAGATAACAGGATAAGCAAATACCCAACCCTCAAATTGTTCGTAATG
GGATGATGATGAAGAGAGAATACAGGGGTCAAGCGATCAGTGAAAGCATTGGCAGATTA

727/237

FIGURE 127

AGAGGCCTCTGGAAAGTTGTCCGGGTGTTGCCGCNGGAGGCCGGTCGAGAGGACNAGG
TGCGCTGCCTGGAGAACCTCCTCCGCTGCCGTCGGCTCCGGAGCCCAGCCCTTCCTAACC
AACCCAACCTAGCCCNGTCCCAGCCGCCAGCGCCTGTCCTGTCNCGGANCCAGCGTNACC
ATGCATCCTGCCGTCTCCTATCCTTACCCGACCTCAGATGCTCCCTCTGCTCCTGGTAAC
TTGGGTTTTACTCCTGTAACAAC TGAAATAACNNGTCTTGATAACNNAGAATATAGATGAAA
TTTAAACNATGCTGATGTGGCTTAGTCATTTTATGCTGACTGGTGTCTTCAGTCAG
ATGTGGCATCCAATTTGAGGANGCTTCCGATGTCATTAAGGAAGAATTCCAAATGAAAA
TCAAGTAGTGTGTTGCCAGAGTTGATTGTGATCAGCACTCTGACATAGCCCAGAGATAACAGGA
TAAGCAAATAACCAACCCCTCAAATTGTTGTAATGGGATGATGATGAAGAGAGAATACAGG
GGTCAGCGATCAGTGAAAGCATTGGCAGATTACATCAGGC

728/237

FIGURE 128

GCCCACGCGTCCG**ATGGCGTT**CACGTTCGCGCCTCTGCTACATGCTGGCGCTGCTCA
CTGCCCGCCTCATCTTCTTCGCCATTGGCACATTATAGCATTGATGAGCTGAAGACTGAT
TACAAGAATCCTATAGACCAGTGTAAATACCCCTGAATCCCCTGTACTCCCAGAGTACCTCAT
CCACGCTTCTTCTGTGCATGTTCTTGCGAGCAGAGTGGCTTACACTGGGTCTCAATA
TGCCCCTCTGGCATATCATATTGGAGGTATATGAGTAGACCAGTGTAGGTGGCCAGGA
CTCTATGACCCCTACAACCATCATGAATGCAGATATTCTAGCATATTGTCAGAAGGAAGGATG
GTGCAAATTAGCTTTATCTTCTAGCATTTCCTACTACCTATATGGCATGATCTATGTT
TGGTGGAGCTCT**TAG**AACAAACACACAGAACATTGGTCCAGTTAAGTGCATGCAAAAAGCCAC
CAAATGAAGGGATTCTATCCAGCAAGATCCTGTCCAAGAGTAGCCTGTGGAATCTGATCAGT
TACTTTAAAAAAATGACTCCTTATTTTAAATGTTCCACATTTGCTTGTGGAAAGACTG
TTTCATATGTTACTCAGATAAAAGATTTAAATGGTATTACGTATAAATTAATATAAAAT
GATTACCTCTGGTGTGACAGGTTGAACTTGCACCTTAAGGAACAGCCATAATCCTCTG
AATGATGCATTAATTACTGACTGTCCTAGTACATTGGAAGCTTGTGTTATAGGAACATTGTA
GGGCTCATTTGGTTCATGAAACAGTATCTAATTATAAATTAGCTGTAGATATCAGGTGC
TTCTGATGAAGTAAAAATGTATATCTGACTAGTGGAAACTTCATGGGTTCTCATCTGTC
ATGTCGATGATTATATGGATACATTACAAAAATAAAAGCGGAAATTCCCTCGCTT
GAATATTATCCCTGTATATTGCATGAATGAGAGATTCCATATTCCATCAGAGTAATAAA
TATACTTGCTTAAATTCTTAAGCATAAGTAAACATGATATAAAATATGCTGAATTACTT
GTGAAGAATGCATTTAAAGCTATTTAAATGTGTTTATTGTAAGACATTACTTATTAAG
AAATTGGTTATTATGCTTACTGTTCTAATCTGGTGGTAAAGGTATTCTTAAGAATTGCAGG
TACTACAGATTTCAAAACTGAATGAGAGAAAATTGTATAACCCTGCTGTTCTTAGT
GCAATACAATAAAACTCTGAAATTAAAGACTC

729/239

FIGURE 129

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA23330
<subunit 1 of 1, 144 aa, 1 stop
<MW: 16699, pI: 5.60, NX(S/T): 0
MAFTFAAFCYMLALLTAALIFFAIWHIIAFDELTDYKNPIDQCNCTLNPLVLPEYLIHAFF
CVMFLCAAELTLGLNMPLLAYHIWRYMSRPVMSGPGLYDPTTIMNADILAYCQKEGWCKLA
FYLLAFFYYLYGMIYVLVSS

Important features:**Signal peptide:**

amino acids 1-20

Type II transmembrane domain:

amino acids 11-31

Other transmembrane domain:

amino acids 57-77 and 123-143

730/237

FIGURE 130

ATTATAGCATTGATGAGCTGAAGACTGATTACAAGATCCTATAGACCAGTGTAAATACCCCTG
AATCCCCTTGACTCCAGAGTACCTCATCCACGCTTCTTGTGTCAATGTTCTTGTC
AGCAGAGTGGCTTACACTGGGTCTCAATATGCCCTCTGGCATATCATATTGGAGGTATA
TGAGTAGACCAGTGTGAGTGGCCAGGACTCTATGACCCCTACAACCACATGAATGCAGAT
ATTCTAGCATATTGTCAGAAGGAAGGATGGTGCAAATTAGCTTTATCTTCTAGCATT
TTACTACCTATATGGCATGATCTATGTTGGTGAGCTCTAGAACACACAGAAGAATT
GGTCCAGTTAAGTGCATGCAAAAGCCACCAATGAAGGGATTCTATCCAGCAAGATCCTGT
CCAAGAGTAGCCTGTGGAATCTGATCAGTTACTTAAAAAATG

131/237

FIGURE 131

CGGACGCGTGGGGAAACCCCTCGAGAAAACAGCAACAAGCTGAGCTGCTGTGACAGAGGG
GAACAAGATGGCGGCCGAAGGGGAGCCTCTGGTGAGGACCCAACTGGGCCTCCGCCGC
TGCTGCTGCTGACCATGCCCTGGCCGGAGGTTGGGACCGCTCGGCTGAAGCATTGAC
TCGGCTTGGGTGATA CGCGTCTGCCACC GGCGCTGT CAGTTGACCTACCCCTTGACAC
CTACCCCTAAGGAAGAGGAGTTGTACGCATGTCAGAGAGGTTGCAGGCTGTTCAATTGTC
AGTTTGTGGATGATGGAATTGACTTAAATCGAACTAAATTGGAATGTGAATCTGCATGTACA
GAAGCATATTCCAATCTGATGAGCAATATGCTGCCATCTGGTGCAGAACATCAGCTGCC
ATTCGCTGAACTGAGACAAGAACAACTTATGTCCTGATGCCAAAAATGCACCTACTCTTC
CTCTAACTCTGGTGAGGT CATTCTGGAGTGACATGATGGACTCCGCACAGAGCTTCATAACC
TCTTCATGGACTTTTATCTTCAAGCCATGACGGAAAAATAGTTATATTCCAGTCTAAGCC
AGAAAATCCAGTACGCACCACATTGGAGCAGGAGCCTACAAATTGAGAGAATCATCTCAA
GCAAAATGTCTATCTGCAAATGAGAAATTACAAGCGCACAGGAATTTCCTGAAGATGGA
GAAAGTGTGGCTTTTAAGATGCCCTCTCTTAACCTCTGGTGAGTTAAGTACAACACTCT
TGT CCTCTCGGTGATGGTATTGCTTGGATTGTTGCAACTGTTGCTACAGCTGGAGC
AGTATGTTCCCTCTGAGAAGCTGAGTATCTATGGT GACTGGAGTTATGAATGAACAAAAG
CTAAACAGATATCCAGCTTCTCTGTGGTTGTTAGATCTAAAAGTGAATCTGCTCATTCTGA
AGCAGGGCCTCACCTACAAAAGTGAATCTGCTCATTCTGA
AAAGACAAGTGTAATAGACATCTAAATTCCACTCCTCATAGAGCTTTAAAATGGTTCA
TTGGATATAGGCCTTAAGAAATCACTATAAAATGCAAATAAGTTACTCAAATCTGTG

732/237

FIGURE 132

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA26847
<subunit 1 of 1, 323 aa, 1 stop
<MW: 36223, pI: 5.06, NX(S/T): 1
MAAPKGSLWVRTQLGLPPLLLTMALAGGSGTASAEAFDSVLGDTASCHRACQLTYPLHTYP
KEEELYACQRGCRLFSICQFVDDGIDLNRTKLECESACTEAYSQSDEQYACHLGCQNQLPFA
ELRQEQLMSLMPKMHLLFPLTLVRSFWSDMMDSAQSFITSSWTFYLQADDGKIVIFQSKPEI
QYAPHLEQEPTNLRESSLSKMSYLQMRNSQAHRNFLEDGESDGFLRCLSLNSGWILTTLVL
SVMVLLWICCATVATAVEQYVPSEKLSIYGDLEFMNEQKLNRYPASSLVVVRSKTEDHEEAG
PLPTKVNLAHSEI

Important features:**Signal peptide:**

amino acids 1-31

Transmembrane domain:

amino acids 241-260

N-glycosylation site.

amino acids 90-93

733 1237

FIGURE 133

TTGGGTGATACGGCGTCTGCCACCGGGCTGTCAGTTGACCTACCCCTGCACACCTACCC
TAAGGAAGAGGAGTTGATCGCATGTCAGAGAGGTTGCAGGCTGTTCAATTGTCAGTTG
TGGATGATGGAATTGACTTAAATCGAACTAAATTGGAATGTAATCTGCATGTACAGAAGCA
TATTCCAATCTGATGAGCAATATGCTTGCCATCTGGTTGCCAGAACATCAGCTGCCATT CGC
TGAAC TGAGACAAGAACAACTTATGTCCTGATGCCAAAATGCACCTACTCTTCCTCTAA
CTCTGGT GAGGT CATTCTGGAGTGACATGATGGACTCCGC

734/237

FIGURE 134

CACACTGGCCGGATCTT TAGAGTCCTTGACCTTGACCAAGGGTCNGGAAAACAGCAACAA
GCTGAGCTGCTGTGACAGAGGAAACAAGATGGCGGCCGAAGGGAGGCCTTGGGTGAGGAC
CCAAGTGGGCTCCCGCCGCTGCTGCTGACCATGGCCTGGCGGAGGTTGGGGACCG
CTTCGGCTGAAGCATTGACTCGGTCTGGGTGATACGGCGCTTGCCACCGGGCTGTCAG
TTGACCTACCCCTTGACACCTACCTAAGGAAGAGGAGTTGTACGCATGTCAGAGAGGTTG
CAGGCTGTTTCAATTGTCAGTTGTGGATGATGGAATTGACTTAAATCGAACTAAATTGG
AATGTGAATCTGCATGTACAGAACATATTCCAATCTGATGAGCAATATGCTGCCATCTT
GGTGCAGAACATCAGCTGCCATTGCTGAAC TGAGACAAGAACAACTTATGTCCTGATGCC
AAAAATGCACCTACTCTTCCTCTAACTCTGGTGAGGTCAATTCTGGAGTGACATGATGGACT
CCGC

735/237

FIGURE 135

GCGAGGTGGCGATCGCTGAGAGGCAGGAGGGCCGAGGCAGGCCTGGGAGGCAGGCCGGAGGT
GGGGCGCCGCTGGGGCCGGCCCGACGGGCTCATCTGAGGGCGCACGGCCCGCAGCGAGC
GTGCGGACTGGCCTCCAAAGCGTGGGGCGACAAGCTGCCGGAGCTGCAATGGGCCGCGGCTG
GGGATTCTTGTGCTGGCCTCCTGGCGCCGTGGCTGCTCAGCTGGGCCACGGAGAGGAGC
AGCCCCCGGAGACAGCGGCACAGAGGTGCTGCCAGGTTAGTGGTTACTTGGATGATTGT
ACCTGTGATGTTGAAACCATTGATAGATTAATAACTACAGGCTTTCCAAGACTACAAAAA
ACTTCTTGAAAGTGACTACTTAGGTATTACAAGGTAAACCTGAAGAGGCCGTGTCCTTCT
GGAATGACATCAGCCAGTGTGAAAGAAGGGACTGTGCTGTCAAACCATGTCATCTGATGAA
GTTCTGATGGAATTAAATCTGCGAGCTACAAGTATTCTGAAGAAGCCAATAATCTCATTGA
AGAATGTGAACAAGCTGAACGACTGGAGCAGTGGATGAATCTCTGAGTGAGGAAACACAGA
AGGCTGTTCTCAGTGGACCAAGCAGTGTGATTCTCAGATAACTCTGTGAAGCTGATGAC
ATTCACTGCCCCTGAAGCTGAATATGTAGATTGCTCTTAATCCTGAGCGCTACACTGGTTA
CAAGGGACCAGATGCTGGAAAATATGGAATGTCACTACGAAGAAAATGTTAAGCCAC
AGACAATTAAAAGACCTTAAATCCTTGGCTCTGGTCAAGGGACAAGTGAAGAGAACACT
TTTACAGTGGCTAGAAGGTCTGTGTAGAAAAAAAGAGCATTCTACAGACTTATATCTGG
CCTACATGCAAGCATTAAATGTGCATTGAGTGCAGAATATCTTACAAGAGACCTGGTTAG
AAAAGAAAATGGGGACACAACATTACAGAATTCAACAGCGATTGATGGAATTGACTGAA
GGAGAAGGTCCAAGAAGGCTTAAGAACTTGTATTTCTACTTAATAGAACTAAGGGCTT
ATCCAAAGTGTACCATTCTCGAGCGCCAGATTCAACTCTTACTGGAATAAAATTC
AGGATGAGGAAAACAAATGTTACTTCTGGAAATACTTCATGAAATCAAGTCATTCTTGTG
CATTTGATGAGAATTCTGATTTGCTGGGATAAAAAGAAGCACACAAACTAAAGGAGGA
CTTCGACTGCATTTAGAAATATTCAAGAATTATGGATTGTGTTGGTTGTTAAATGTC
GTCTGTGGGAAAGCTTCAGACTCAGGGTTGGCACTGCTCTGAAGATCTTATTTCTGAG
AAATTGATAGCAAATATGCCAGAAAGTGGACCTAGTTATGAATTCCATCTAACAGACAAGA
AATAGTATCATTATTCAACGCATTGGAAGAATTCTACAAGTGTGAAAGAATTAGAAAATC
TCAGGAACCTGTTACAGAATATTCATAAAGAAAACAAGCTGATATGTGCCTGTTCTGGAC
AATGGAGGCAGAAAGAGTGGATTTCATTCAAAGGCATAATAGCAATGACAGTCTAACCAA
ACATTATATAAAAGTGCTTTGTAAAGGAGAATTATATTGTTAAAGTAAACACATT
AAAAATTGTGTTAAAGTCTATGTATAACTACTGTGAGTAAAGTAATACTTTAATAATGTG
GTACAAATTAAAGTTAATATTGAATAAAAGGAGGATTATCAAATTAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAA

736(237)

FIGURE 136

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53974
<subunit 1 of 1, 468 aa, 1 stop
<MW: 54393, pI: 5.63, NX(S/T): 2
MGRGWGFLFGLLGAVWLLSSGHGEEQPPETAAQRCFCQVSGYLDCTCDVETIDRFNNYRLF
PRLQLLLESFYKVNLRPCPFWNNDISQCGRDCAVKPCQSDEVPDGIKSASYKYSEEA
NNLIEECEQAERLGAVDESLSEETQKAVLQWTKHDDSSDNFCEADDIQSPEAEYVDLLLNP
RYTGYKGPDWKIWNVIYEENCFKPQTIKRPLNPLASGQGTSEENTFYSWLEGLCVEKRAFY
RLISGLHASINVHLSARYLLQETWLEKKWGNITEFQQRFDGILTEGEGPRLKNLYFLYLI
ELRALSKVLPFFERPDFQLFTGNKIQDEENKMLLEILHEIKSFPLHFDENSFFAGDKKEAH
KLKEDFRLHFRNISRIMDCVGCFKCRLWGKLQTQGLGTALKILFSEKLIANMPESGPSYEFH
LTRQEIVSLFNAFGRISTSVKELENFRNLLQNIH
```

Important features:**Signal peptide:**

amino acids 1-23

N-glycosylation site.

amino acids 280-283 and 384-387

Amidation site.

amino acids 94-97

Glycosaminoglycan attachment site.

amino acids 20-23 and 223-226

Aminotransferases class-V pyridoxal-phosphate

amino acids 216-222

Interleukin-7 proteins

amino acids 338-343

737/237

FIGURE 137

GCTGGAAATATGGATGTCATCTACGAGAACTGTTTAAGCCACAGACAATTAAAAGACCTT
TAAATCCTTGGCTTCTGGTCAAGGGACAAGTGAAGAGNACACTTTTACAGTGGCTAGAA
GGTCTCTGTAGAAAAAAGAGCATTCTACAGACTTATATCTGGCCTACATGCAAGCATTAA
TGTGCATTTGAGTGCAAGATATCTTTACAAGAGACCTGGTAGAAAAGAAATGGGGACACA
ACATTACAGAATTNAACAGCGATTTGATGGAATTTGACTGAAGGAGAAGGTCCAAGAAGG
CTTAAGAACTTGTATTTCTACTTAATAGAACTAAGGGCTTATCAAAGTGTACCATT
CTTNGAGGCCAGATTTCAACTNTTACTGGAAATAAATTCAAGGATGAGGNAACAAAA
TGTTACTTTGGAAATACCTCATGAAATCAAGTCATTCCTTGCATTTGATGAGAATTCA
TTTTTTGCTG

738/237

FIGURE 138

CGGACGCGTGGCGGACGCGTGGCGGACGCGTGGTTGGGAGGGGGCAGGATGGGAGGGAA
AGTGAAGAAAACAGAAAAGGAGAGGGACAGAGGCCAGAGGACTCTCATACTGGACAGAAC
CGATCAGGCATGGAACCTCCCTCGTCACTCACCTGTTCTGCCCTGGTCTCCTGACAGG
TCTCTGCTCCCCCTTAACCTGGATGAACATCACCCACGCCATTCCAGGGCCACCAGAAG
CTGAATTGGATACAGTGTCTTACAACATGTTGGGGTGGACAGCGATGGATGCTGGTGGC
GCCCTGGATGGGCCTTCAGGCACCGGAGGGGGACGTTATCGCTGCCCTGTAGGGGG
GGCCCACAATGCCCATGTGCCAAGGCCACTTAGGTGACTACCAACTGGAAATTCATCTC
ATCCTGCTGTGAATATGCACCTGGGATGTCCTGTTAGAGACAGATGGTATGGGGATTC
ATGGTGAGCTAAGGAGAGGGTGGCAGTGTCTGAAGGCCATAAAAGAAAAAGAGAA
GTGTGGTAAGGGAAAATGGTCTGTGTGGAGGGTCAAGGAGTTAAAACCCTAGAAAGCAA
AGGTAGGTAATGTCAGGGAGTAGTCTCATGCCCTCAACTGGGAGCATGTTCTGAGGGT
GCCCTCCAAGCCTGGAGTAACTATTCCTCCATCCCCAGGCCTGTGCCCTCTGGTCT
CGTGCTTGTGGCAGCTGTCTCAGTTCTGGATATGTGCCGTGTGGATGCTTCATTCCA
GCCTCAGGGAAAGCCTGGCACCCACTGCCAACGTGAGCCAGAGGAAGGCTGAGTACTTGGTT
CCCAGAAGGAGATACTGGGTGGAAAAAGATGGGCAAAGCGGTATGATGCCCTGGCAAAGGG
CCTGCATGGCTATCCTCATTGCTACCTAACGTGCTGCTGAAAGCTCCATGTTCTAACAGA
TTCAGACTCCTGGCCAGGTGTGGCCACACCTGTAATTCTAGCAGTTGGGAGGCCAAG
GTGGGCAGATCACTTGAGGTCAAGACCAGCCTGCCAACATGGTAAACTCCAT
CTCTACTAAAAAAAAAAACAAAAATTAGCTGGGTGCGCTAGTCATGCCGTAAATCTC
ATCTACTCGGGAGGCTAACGACAGGAGACTCTCACTCAACCCAGGAGGTGGAGGGTGCCTG
AGCCAAAGATTGTGCCTCTGCACTCTAGCGTGGGTGACAGAGTAAGCGAGACTCCATCTCAA
AATAATAATAATAATTCAAGACTCCTTATCAGGAGTCCATGATCTGGCCTGGCACAGTAA
CTCATGCCGTAAATCCAAACATTGGGAGGCCAACGCAGGAGGATTGCTTGAGGTCTGGAG
GTTTGAGACCAGCCTGGCAACATAGAAAGACCCATCTCTAAATAATGTTTAAAAAT

139/237

FIGURE 139

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57039
><subunit 1 of 1, 124 aa, 1 stop
><MW: 13352, pI: 5.99, NX(S/T): 1
MELPFVTHLFLPLVFLTGLCSPFNLDHHPRLFPGPPEAEFGYSVLQHVGGGQRWMLVGAPW
DGPSGDRRGDVYRCPVGGAHNAPCAKGHLGDYQLGNSSHPAVNMHLGMSLLETDGDGGFMVS
```

Important features:**Signal peptide:**

amino acids 1-22

Cell attachment sequence.

amino acids 70-73

N-glycosylation site.

amino acids 98-101

Integrins alpha chain proteins

amino acids 67-81

740/237

FIGURE 140

CACAGTTCCCCACCATCACTCNTCCCATTCTTCCAACTTTATTTAGCTTGCCTTGGGA
GGGGCAGGATGGGAGGGAAAGTGAAGAAAACAGAAAAGGGAGAGGGACAGAGGCCAGAGGAC
TTCTCATACTGGACAGAAACCGATCAGGCATGGAACCTCCCTCGTCACTCACCTGTTCTTG
CCCCTGGTGTCTGACAGGTCTCTGCTCCCCCTTAACCTGGATGAACATCACCCACGCCT
ATTCCCAGGGCCACCAGAAGCTGAATTGGATACAGTGTCTAACACATGTTGGGGTGGAC
AGCGATGGATGCTGGTGGCGCCCCCTGGGATGGGCTTCAGGCGACCGGAGGGGGACGTT
TATCGCTGCCCTGTAGGGGGGCCACAATGCCCATGTGCCAAGGGCCACTTAGGTGACTA
CCAACCTGGAAATTCTCATCCTGCTGTGAATATGCACCTGGGATGTCTGTTAGAGA
CAGATGGTGATGG

7471237

FIGURE 141

AAAGTTACATTTCTCTGGAACCTCCTAGGCCACTCCCTGCTGATGCAACATCTGGGTTTG
GGCAGAAAGGAGGGTGCTCGGAGCCGCCCTTCTGAGCTTCCTGGGCCGGCTCTAGAACAA
ATTCAAGGCTTCGCTGCGACTCAGACCTCAGCTCCAACATATGCATTCTGAAGAAAGATGGCT
GAGATGGACAGAAATGCTTATTTGAAAGAAACAATGTTCTAGGTCAAACACTGAGCTACCA
AATGCAGACTTCACAATGTTCTAGAAGAAATCTGGACAAGTCTTTCATGTGGTTTTCT
ACGCATTGATTCCATGTTGCTCACAGATGAAGTGGCCATTCTGCCTGCCCTCAGAACCTC
TCTGTACTCTCAACCAACATGAAGCATCTCTGATGTGGAGGCCAGTGATCGCGCTGGAGA
AACAGTGTACTATTCTGTCGAATACACAGGGGAGTACGAGAGGCCGTACACGAGCCACATCT
GGATCCCCAGCAGCTGGTGCCTACTCACTGAAGGTCTGAGTGTGATGTCACTGATGACATC
ACGGCCACTGTGCCATACAAACCTTCGTGTCAGGGCCACATTGGGCTCACAGACCTCAGCCTG
GAGCATCCTGAAGCATCCCTTAATAGAAACTCAACCACCTTACCCGACCTGGGATGGAGA
TCACCAAAAGATGGCTTCCACCTGGTTATTGAGGCTGGAGGACCTGGGCCAGTTGAGTTC
CTTGTGGCCTACTGGAGGGAGGCTGGTGCAGGAAACATGTCAAAATGGTGGAGGAGTGG
GGGTATTCCAGTGCACCTAGAAACCATGGAGCCAGGGCTGCATACTGTGTGAAGGCCAGA
CATTCGTGAAGGCCATTGGGAGGTACAGCGCCTTCAGCCAGACAGAAATGTGTGGAGGTGCAA
GGAGAGGCCATTCCCCTGGTACTGGCCCTGTTGCCTTGTGGCTTCATGCTGATCCTTGT
GGTGTGCCACTGTTCGTCTGGAAAATGGCCGGCTGCTCCAGTACTCCTGTTGCCCGTGG
TGGTCCCTCCAGACACCTGAAAATAACCAATTACCCCCAGAAGTTAACAGCTGCAGAAGG
GAGGAGGGTGGATGCCACGGCTGTGATGTCTCTGAGGAACCTCAGGGCTGGAT
CTCAT**TAG**TTGCGGAAGGCCAGGTGAAGCCGAGAACCTGGTCTGCATGACATGGAAACC
ATGAGGGGACAAGTTGTGTTCTGTTCTGTTCCGCCACGGACAAGGGATGAGAGAAGTAGGAAGA
GCCTGTTGTCTACAAGTCTAGAACGCAACCACAGAGGCAAGGGTGGTTGTCTAACAGAACAC
TGACTGAGGCTTAGGGGATGTGACCTCTAGACTGGGGCTGCCACTGCTGGCTGAGCAACC
CTGGAAAAGTGACTTCATCCCTCGGTCTAAGTTCTCATCTGTAATGGGGAAATTACC
TACACACCTGCTAACACACACACAGAGTCTCTCTATATATACACACGTACACATAAA
TACACCCAGCACTGCAAGGCTAGAGGGAAACTGGTGACACTCTACAGTCTGACTGATTGAG
TGTTCTGGAGAGCAGGACATAATGTATGATGAGAATGATCAAGGACTCTACACACTGGGT
GGCTTGGAGAGGCCACTTCCCAGAATAATCCTTGAGAGAAAAGGAATCATGGAGCAATGG
TGTTGAGTTCACTTCAAGCCAATGCCGGTGCAGAGGGGAATGGCTTAGCGAGCTCTACAGT
AGGTGACCTGGAGGAAGGTACAGCCACACTGAAAATGGGATGTGCTGATGAAACACGGAGGATC
CATGAACACTGTAAAGTGTGACAGTGTGTCACACTGCAGACAGCAGGTGAAATGTATGT
GTGCAATGCGACGAGAATGCAAGTCAGTAACATGTGCTGATGTTGTGCTCCTTTTC
TGTTGGTAAAGTACAGAATTCAAGAAATAAAAGGCCACCCCTGGCCAAAAGCGGTAAAAAA
AAAAAAAAAA

742 1237

FIGURE 142

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57033
<subunit 1 of 1, 311 aa, 1 stop
<MW: 35076, pI: 5.04, NX(S/T): 2
MQTFTMVLEEIWTSLFMWFYALIPCLLTDEVAILPAPQNLSQLSTNMKHLLMWSPVIAPGE
TVYYSVQEYQGEYESLYTSHIWIPSSWCSLTEGPECDVTDDITATVPYNLRVRATLGSQTSAW
SILKHPFNRNSTILTRPGMEITKDGFLVIELEDLGPQFEFLVAYWRREPGAEEHVKMVRSG
GIPVHLETMEPGAAYCVKAQTFVKAIGRYSAFSQTECVEVQGEAIPLVLALFAFVGFMLILV
VVPLFVWKMGRLLQYSCCPVVLPDTLKINTSPQKLISCRREEVDACATAVMSPEELLRAWIS
```

Important features:**Signal peptide:**

amino acids 1-29

Transmembrane domain:

amino acids 230-255

N-glycosylation site.

amino acids 40-43 and 134-137

Tissue factor proteins.

amino acids 92-119

Integrins alpha chain proteins

amino acids 232-262

143/237

FIGURE 143

TCCTGCTGATGCACATCTGGTTGGCAAAAGGAGGTTGCTCGAGCCGCCCTTAGCTT
CCTGGCCGGCTCTAGAACATTAGGCTCGCTGCGACTAGACCTCAGCTCCAACATATGCA
TTCTGAAGAAAGATGGCTGAGATGACAGAAATGCTTATTTGGAAAGAAACAATGTTCTAGG
TCAAACGTGAGTCTACCAAATGCAGACTTCACAATGGTTCTAGAAGAAATCTGGACAAGTCT
TTCATGTGGTTTCTACGATTGATTCCATGTTGCTCACAGATGAAGTGGCCATTCTGC
CTGCCCTCAGAACCTCTCTGTACTCTAACCAACATGAAGCATCTCTGATGTGGAGCCA
GTGATCGCGCCTGGAGAAACAGTGTACTATTCTGTCGAATACCAGGGGGAGTACGAGAGCCT
GTACACGAGCCACATCTGGATCCCCAGCAGCTGGTGTCACTCACTGAAGGTCTGAGTGTG
ATGTCACTGATGACATCACGGCCACTGTGCCATACAACCTTGTGTCAGGGCACATTGGC
TCACAGACCTCAGCCTGGAGCATCCTGAAGCATCCCTTAATAGAAACTCAACCACCTTAC
CCGACCTGGGATGGAGATCACCAAAGATGGCTNCACCTGGTTATTGAGCTGGAGGACCTGG
GGCCCCAGTTGAGTTCCCTGTGGCCTANTGGAGGAGGGCGAACCCCTGCGCGCAAGGG
GTTNGCGAACCCCTTGCGGCCGCTGGGTATCTCTCGAGAAAAGAGAGGCCAATATGACCC
ACATACTCAATATGGACGAANTGCTATTGTCCACCTGTTGAGTGGCGCTGGTTGAT

744 | 237

FIGURE 144

CCACCGCGTCCGCCACCGTCCGAGGGACAAGAGAGAAGAGAGACTGAAACAGGGAGAAGA
GGCAGGAGAGGAGGAGGTGGGGAGAGCACGAAGCTGGAGGCCACACTGAGGGAGGGCGGGA
GGAGGTGAAGAAGGAGAGAGGGGAGAAGAGGCAGGGAGCTGGAAAGGAGAGAGGGAGGAGGAG
GAGGAGATGCCGGATGGAGACCTGGAGTTAGGTGGCTTGGGAGAGCTTAATGAAAAGAGAAC
GGAGAGGAGGTGTGGGTTAGGAACCAAGAGGTAGCCCTGTGGCAGCAGAAGGCTGAGAGGA
GTAGGAAGATCAGGAGCTAGAGGGAGACTGGAGGGTCCGGAAAAGAGCAGAGGAAAGAGG
AAAGACACAGAGAGACGGGAGAGAGAAGAAGAGTGGGTTGAAGGGCGGATCTCAGTCCTG
GCTGCTTGGCATTGGGAACTGGACTCCCTGTGGGAGGGAGAGGAAAGCTGGAAAGTC
GGAGGGACAGGGTCCCAGAAGGAGGGACAGAGGAGCTGAGAGAGGGGGCAGGGCGTTGGG
CAGGGTCCCTCGGAGGCCTGGGGATGGGGCTGCAGCTCGTCTGAGCGCCCTCGAGC
GCTGGTACTCTGGGCTGCACTGGGGCAGCAGCTCACATCGGACCAGCACCTGACCCGAGG
ACTGGTGGAGCTACAAGGATAATCTCAGGGAAACTCGTGTGCCAGGGCTCCTTCTGGG
CTGGTGAATGCAGCGTGGAGTCTGTGTGCCAGGGAGCAGAGCCCCGTGGATGTGGA
GCTGAAGAGGGTTCTTATGACCCCTTCTGCCCTTAAGGCTAGCACTGGAGGAGAGA
AGCTCCGGGAAACCTTGACACACCGGCCACATGTCTCCTCGCCTGCACCCGACCT
GTGGTCAATGTCTGGAGGTCCCTCCTTACAGCCACCGACTCAGTGAAC TGCGGCTGCT
GTTTGGAGCTCGCGACGGAGCCGGCTCGAACATCAGATCAACCAACCAGGGCTCTGCTG
AGGTGCAGCTATTCAACCAAGGAACCTACGGGAATTCAAGCGCTGCCTCCCGGGC
CCCAATGCCCTGGCATTCTCAGCCTTGTCAACGTTGCCAGTACCTCTAACCCATTCT
CAGTCGCCTCTTAACCGCGACACCATCACTCGCATCTCCTACAAGAATGATGCCTACTTC
TTCAAGACCTGAGCCTGGAGCTCCTGTTCCCTGAATCCTCGGCTTCATCACCTATCAGGGC
TCTCTCAGCACCCGCCCTGCTCCGAGACTGTCACCTGGATCCTCATTGACCGGGCCCTCAA
TATCACCTCCCTCAGATGCACTCCCTGAGACTCCTGAGCCAGAACCTCCATCTCAGATCT
TCCAGAGCCTCAGCGGTAAACAGCCGCCCTGCAGCCCTGGCCACAGGGCACTGAGGGG
AACAGGGACCCCGGCACCCCGAGAGGGCGTGCAGGCCCCAACTACCGCCTGCATGTGGA
TGGTGTCCCCCATGGTCGCTTGAACTCCCTCGAGGATTGCACCCGCCGTCTAACCT
CCCACAAGGCAGGGAGTTACCCCTAAAACAAAGCTATTAAAGGGACAGAATACTTA

145/237

FIGURE 145

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA34353
<subunit 1 of 1, 328 aa, 1 stop
<MW: 36238, pI: 9.90, NX(S/T): 3
MGAAARLSAPRALVLWAALGAAAHHGPAPDPEDWWWSYKDNLQGNFVPGPPFWGLVNAAWSLC
AVGKRQSPVDVELKRVLYDPFLPPLRLSTGGEKLRGTLNTGRHVSFLPAPRPVVNVSGGPL
LYSHRLSELRLLFGARDGAGSEHQINHQGFSAEVQLIHFNQELYGNFSAASRGPNGLAILSL
FVNVASTSNPFLSRLLNRTDTITRISYKNDAYFLQDLSLELLFPESFGFITYQGSILSTPPCSE
TVTWILIDRALNITSLQMHSRLLSQNPPSQIFQSLSGNSRPLQPLAHRALRGNRDPRHPER
RCRGPNYRLHVVDGVPHGR

Important features:**Signal peptide:**

amino acids 1-23

Transmembrane domain:

amino acids 177-199

N-glycosylation site.

amino acids 118-121, 170-173 and 260-263

Eukaryotic-type carbonic anhydrases proteins

amino acids 222-270, 128-164 and 45-92

746/237

FIGURE 146

GGCGCCTGGTTCTGCGGTACTGGCTGTACGGAGCAGGAGCAAGAGGTCCGCCAGCCTCCGCCAGCCT
GTTCTGTCCCCCCCCTCGCTCCGTCACTGCAGCTACTGCTCAGAAACGCTGGGGGCCACCCCTGGCAGACTAACGAA
GCAGCTCCCTTCCCACCCAACTGCAGGTCAATTGGACGCTTGCCATTCTCCAGGTTGAGGGAGC
CGCAGAGGCGGAGGCTCGGTATTCTGCAGTCAGCACCCACGTGCCCGGACGCTCGGTGCTCAGGCCCTC
GCCAGCGGGCTCTCGCTCGGTCCCTGTGAAGGCTCTGGCGGCTGCAGAGGCCGCGTCCGGTTGGCT
CACCTCTCCAGGAAACTAACATGGAGGCCAAGGGAGTGGAGAGGCCCTGTCTTGGAGATTTCTGGGAA
ATCCTGAGGTATTCAATTATGAGGTACCGCGGGAGTGGCTCAGAGTAACCACAGTGTGTTCATGGCTAGA
GCAATTCCAGCCATGGTGGTCCCAATGCCACTTTATTGGAGAAACTTTGGAAAATACATGGATGAGGATGGT
GAGTGGTGGATAGCAAACAACAGAGGAAAGGGCCATCACAGACAATGACATGCAGAGTATTGGACCTTCAT
AATAAATTACGAAGTCAGGTATCCAACAGCCTCTAATATGGAGTATATGACATGGGATGTAGAGCTGGAAAGA
TCTGCAGAATCTGGGCTGAAAGTTGCTTGTGGGAAACATGGACCTGCAAGCTGCTTCCATCAATTGGACAGAA
TTGGGAGCAGACTGGGGAGATATAGGCCCGACGTTCATGTACAATCGTGGTATGATGAAGTGAAGACTTT
AGCTACCCATATGAACATGCAACCCATATTGTCATTCAAGGTGTTCTGCCCTGTATGTCACATTTACA
CAGGTGTTGGCAACTAGAACAGAACAGTGGTGTGCAATTACTCCCAAAGGGAAACTGGTGGGCCATGCCCTACAAA
CATGGCGGCCCTGTTCTGCCACCTAGTTGGAGGGGCTGTAGAGAAAATCTGTGCTACAAAGAAGGG
TCAGACAGGTATTATCCCCCTGAGAAGAGGAAACAAATGAAATAGAACGACAGCAGTCACAAGTCCATGACACC
CATGTCGGACAAGATCAGATGATAGTAGCAGAAATGAAGTCATAAGGCACAGCAAATGCCAAATTGTTCT
TGTGAAGTAAGATTAAGAGATCAGTGCAAAGAACCTGCAATAGGTACGAATGTCCTGCTGGCTGTTGGAT
AGTAAAGCTAAAGTTGGCAGTGTACATTATGAAATGCAATTCCAGCATCTGTAGAGCTGCAATTCTGTT
ATAATAGACAATGATGGTGGCTGGTAGATATCACTAGACAAGGAAGAACGATTATTCATCAAGTCAAA
AATGGTATTCAAAACTTGGCAATATCAGTGTGTAATTCTTCAGTCTCTAAAGTAAACAGTCAGGCTG
ACTTGTGAAACAACAGTGGACAGCTCTGTCATTCAAGCCTGTTCACATTGCCAAGAGTATACTGTCCT
CGTAACGTATGCAAGCAAATCCACATTATGCTGTAATTGAAACTCGAGTTATTCTGATCTGCAAGTAC
TGCAGAGCAGCAGTACATGCTGGAGTGGTTGAAATCACGGTGGTTATGTGATGTAATGCCCTGGACAAAAGA
AAGACACTACATTGCTTTCAGAAATGGAATCTCTCAGAAAGTTACAGAACCTCCAGGAGGAAAGGCATTC
AGAGTGGTGTGTTGCTGTTGCTGTTGCTGTTGCTGTTGCTGTTGCTGTTGCTGTTGCTGTTGCTGTTGCT
ATTTGTTATAAAACTGTAACATTACTGTACAGACTACATCAACTATTTCAGCCAAAAGGTGCCAAATGCA
TAAATCTGATAAAACAAAGTCTATAAAATAAAATGGGACATTAGCTTGGGAAAGTAAATGAAAATATAATGG
TTTAGAAATCCTGTGTTAAATATTGCTATATTCTTAGCAGTTATTCTACAGTTAAACATAGTCATGATT
GTTCTACGTTCATATATTATGGTCTTGTATATGCCACTAAATGGAAATCTAAACATTGAATGTGAATG
GCCCTCAGAAAATCATCTAGTCATTAAAAATAATCAGACTCTAAAACGTTATCACATTTC
AGTTCAATGCTATGCCATTACCAACTCCAAATAACTCAAATAATTTCACCTTAATAACTGTAAGTTTTC
TGTTAATTAGGCATATAGAATTTAAATTCTGATATTGCACTCTTATTATATAAAATCCTTTAATATC
CAAATGAAATCTGTTAAAATGGTGTGATTCTGGGAAATGGCCTTAAAATGAAATGAAATGCAAGGTGG
ATGAAAACATTCTAGTGTGATGTAATGTAAGGTTAAGCATGGACAGCCAGAGCTTCTATGTA
AAATTGAGGTACATATTCTTTGTATCCTGGCAAATACTCCTGCAGGCCAGGAAGTATAATGCAAAAAGTT
GAACAAAGATGAACATAATGTTAGTACATTACCAATTGCCACTGATTTTTAAATGGTAAATGACCTGTATATAA
ATATTGCCATATCATGGTACCTATAATGGTGTATATTGTTCTATGAAAGGAAATGTTGCTGTTGACTAAA
AATCTGTTAAATGTTAGTTGGTAAATTCTGCTGGGATTTACATATTAAATTTCTGCTGGGATTT
TAAACATTAAAATTAATCATGTTCAAAAAAAAAAAAA

187/237

FIGURE 147

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45417
<subunit 1 of 1, 500 aa, 1 stop
<MW: 56888, pI: 8.53, NX(S/T): 2
MKCTAREWLRTTVLFMARAIPAMVVPNATLLEKLYMDEDGEWWIAKQRGKRAITDNDM
QSILDLNKLRSQVYPTASNMEYMTWDVELERSAESCLWEHGPASLLPSIGQNLGAHW
GRYRPPTFHVQSWYDEVKDFSYPYEHECNPYCPFRCSPVCTHYTQVVWATSNRIGCAINLC
HNMMNIWGQIWPKAVYLVCNYSPKGNWWGHAPYKHGRPCSACPPSFGGGCRENLCYKEGSDRY
YPPPREEETNEIERQQSQVHDTHRTRSDSSRNEVISAQQMSQIVSCEVRLRDQCKGTTCNR
YECPAGCLDSKAKVIGSVHYEMQSSICRAAIHYGIIDNDGGWVDITRQGRKHYFIKSNRNGI
QTIGKYQSANSFTVSKVTQAVTCETTVEQLCPFHKPASHCPRVYCPRNMQANPHYARVIG
TRVYSDLSSICRAAVHAGVVRNHGGYVDVMPVDKRTYIASFQNGIFSESLQNPPGGKAFRV
FAVV

Important features:**Signal peptide:**

amino acids 1-20

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7 protein

amino acids 165-186, 196-218, 134-146, 96-108 and 58-77

N-glycosylation site

amino acids 28-31

748/237

FIGURE 148

GGGGAGACAAGCGCAGAGCGCAGCGCACGGCCACAGACAGCCCTGGCATCCACCGACGGCG
CAGCCGGAGCCAGCAGAGCCGAAGGCAGCGCCCCGGCAGAGAAAGCCGAGCAGAGCTGGT
GGCGTCTCCGGGCCGCGCTCCGACGGGCAGCGCCCTCCCCATGTCCCTGCTCCCACGCCG
CGCCCTCCGGTCAGCATGAGGCTCCTGGCGGCCGCTGCTCCTGCTGCTGCTGGCGCTGT
ACACCGCGCGTGTGGACGGTCAAATGCAAGTGCTCCGGAAGGGACCCAAGATCCGCTAC
AGGCACGTGAAGAAGCTGGAAATGAAGCCTAACAGTACCCGACTGCGAGGGAGAAGATGGTTAT
CATCACCAAGAGCGTGTCCAGGTACCGAGGTAGGAGCACTGCCTGCACCCCAAGCTGC
AGAGCACCAAGCGCTTCATCAAGTGGTACAACGCCCTGGAACGAGAAGCGCAGGGTCTACGAA
GAATAGGGTAAAAACCTCAGAAGGGAAAATCCAAACCAGTTGGGAGACTTGTGCAAAGGA
CTTGCAGATTAAAGCCTTTC
TTTCTCACAGGCATAAGACACAAATTATATTGTTATGAAGCACTTTACCAACGGTCAG
TTTTACATTTATAGCTGCGTGCAGGCTTCCAGATGGGAGACCCATCTCTTGTGCT
CCAGACTTCATCACAGGCTGCTTTATCAAAAAGGGAAAATCATGCCTTCCCTTTAA
AAAATGCTTTGTATTGTCATACGTCACTATACATCTGAGCTTATAAGCGCCCGGA
GGAACAATGAGCTTGGTGGACACATTTCATTGCACTGTTGCTCCATTCTAGCTTGGGAAGC
TTCCGCTTAGAGGTCTGGCGCTCGCACAGCTGCCACGGCTCTGGCTATGGCG
GTCACAGCCTCAGTGTGACTCCACAGTGGCCCTGTAGCCGGCAAGCAGGAGCAGGTCTCT
CTGCATCTGTTCTTGAGGAACACTCAAGTTGGTGCAGAAAATGTGCTTCATTCCCCCT
GGTAATTTCACACACCTAGGAAACATTCCAAGATCCTGTGATGGCGAGACAAATGATC
CTTAAAGAAGGTGTGGGTCTTCCAACCTGAGGATTCTGAAAGGTTCACAGGTTCAATA
TTTAATGCTTCAGAAGCATGTGAGGTTCCAACACTGTCAAGCAAAACCTAGGAGAAAATC
AAAAATATATGAATAACATGCGCAATACACAGCTACAGACACACATTCTGTTGACAAGGGAA
AACCTCAAAGCATGTTCTTCCCTCACCACAAACAGAACATGCAGTACTAAAGCAATATAT
TTGTGATTCCCCATGTAATTCTCAATGTTAAACAGTGCAGTCCTTCTGAAAGCTAAGAT
GACCATGCGCCCTTCCCTGTACATATAACCTTAAGAACGCCCCCTCCACACACTGCC
CAGTATATGCCGCATTGACTGCTGTGTTATGCTATGTACATGTCAAGAAACCATTAGCAT
TGCATGCAGGTTCATATTCTTAAGATGGAAAGTAATAAAATATTTGAAATGTAAAA
AAAAAAAAAA

749/237

WO 00/53756

FIGURE 149

MSLLPRRAPPVSMRLAALLLLLALYTARVDGSKCKCSRKGPKIRYSDVKKLEMKPYPH
CEEKMVIIITTKSVSRYRGQEHLHPKLQSTKRFIKWYNNAWNEKRRVYEE

Signal sequence:
amino acids 1-34

750/237

FIGURE 150

GCCCCAGGGACTGCTATGGCTCCTTGTTCACCCGGTCTGCCTCAGTTAAACTCCAATGTCCTCCTGTG
 GTTAAGTGCCTTGCCATCAAGTTCACCCCTATTGACAGCCAAGCACAGTATCCAGTTGTCACACAAATTATGG
 CAAAATCGGGGCCCTAAGAACACCGTTACCCATGAGATCTTGGGTCCAGTGGAGCAGTACTTAGGGTCCCCTA
 TGCCCTACCCCCACTGGAGAGAGGGGTTTCAAGCCCCAGAACCCCCGTCCTGGCATGCCATCGAAATAC
 TACTCAGTTGCTGCTGTGCCCCAGCACCTGGATGAGAGATCTTACTGATGACATGCTGCCATCTGGTT
 TACCGCCAATTGGATACTTGTGACCATGTTCAAGATCAAATGAAGACTGCCTTACTTAAACATCTACGT
 TACCGCCAATTGGATACTTGTGACCATGTTCAAGATCAAATGAAGACTGCCTTACTTAAACATCTACGT
 GCCCAGGAAAGATGGAGCCAACACAAAGAAAAACGAGATATAACGAGTAATGACCGTGGTGAAGACGAAGA
 TATTGATGATGAGAAGCTCATGGTCTATATCCATGGGGATCTTACATGGAGGGCACCGGCAA
 CATGATTGACGGCAGCATTTGGCAAGCTACGGAAACGTCATCGTGTACACCATTAAACTACCGTCTGGGAATACT
 AGGGTTTAAAGTACCGGTGACCGAGCAGCAAAGGCAACTATGGCTCCTGGATCAGATTCAAGCAGTGCCTG
 GATTGAGGAGAATGTGGAGCCTTGGGGGGACCCAAGAGAGTGAACATCTTGGCTGGGGCTGGGGCTC
 CTGTTGTCAGCCTGTTGACCCCTGCCCCTACTCAGAAGGTCTCTCCAGAAGGGCATATTGAGCAGGGCACCG
 CCTGTCAGCTGGGCACTGAACATACAGCCGGCAAGTACACTCGGATATTGGCAGACAAGGTGGCTGCAACAT
 GCTGGACACCACGGACATGGTAAATGCTGGGAAACAAGAAACTACAAGGAGCTCATCCAGCAGACCATACCCC
 GGCCACCTACCATGCTGGGGGGTGTGACGGCGACGTCATCCAGACGACCCCCAGATCCTGATGGA
 GCAAGCGAGTCCCTCAACTACGACATCATGCTGGCGTCAACCAAGGGGAAGGGCTGAAGTCGTGGACGGCAT
 CGTGGATAACGAGGACGGTGTGACGCCAACGACTTTGACTTCTCCGTGTCACCTCGTGGACAACTTACGG
 CTACCTGAAGGAAAGACACTTGGGGAGACTATCAAGTTGATGACACAGACTGGGCCATAAGGAAACCC
 GGAGACGCGGGAAAACCTGGTGGCTCTTTACTGACCAAGCTGGGCCCCCGTGGCCGGCACCT
 GCACGCGCAGTACGGCTCCACCTACTTCTATGCCCTCATCTACTGCCAAAGGCAAGGGCTGAAAGTCGTGGACGGCTG
 GGCAGATTGGCCATGGTGTGAGGTTCCCTATGCTCTGGCATCCCCATGATCGGTCACCGAGCTCTCAG
 TTGTAACCTTCAAGAACGACGTCATGCTCAGCCGGTGTGACCTACTGGACGAACCTGCCAAAATGG
 TGATCAAATCAACCAAGTCTCAGGATACCAAGTCAACACAAAACCAACCGCTTGAAGAAGTGGCTG
 GTCCAAGTATAATCCAAAGGACGCTATCTGCATATTGGCTGAAACCCAGAGTGAAGAGATCACTACGGGC
 AACGAAAGTGGCTTCTGGTGAACCTGTTCTCATTTGACAACTTGAACGAGATATTCCAGTATGTTCAAC
 AACACAAAGGTTCTCCACCAAGACATGACATCATTCCCTATGGCACCCGGCATCTCCGCAAGATATGGCC
 AACACCAAAGCCCAGCAATCACTCTGCCAACATCCAAACACTCTAAGGACCCCTCACAAACAGGGCTGA
 GGACACAACGTCTCATTGAAACCAAACGAGATTATTCACCGAATTAAAGTGTACCATGGCTGTCGGGGCTG
 GCTCTCTCCTCAACATCTTAGTTGCGGCGTGTACTACAAAAGGACAAGAGGGGCCATGAGACTCACAG
 GCGCCCGAGTCCCCAGAGAACACCAAAATGATATCGCTCACATCCAGAACGAGAGATCATGCTCTGAGAT
 GAAGCAGCTGGAACACGATCACGAGTGTGAGTCGCTGCAGGACACGACACTGAGGCTCACCTGCCGCCAGA
 CTACACCTCACGCTGGGGTGGCTGAGATCACATCCCCTATGACGCCAACACCATCACCAGTATTCCAAA
 CACACTGACGGGGATGAGCTGGCTTGCACACTTTAACACCTTCAGTGGAGGACAAAAGTACAATTTACCCA
 CGGACATTCCACCACTAGAGTATAGCTTGCCTTATTCCTCTATCCCTGCCCTACCCGCTCAGCAACAT
 AGAAGAGGAAAGGAAAG
 GACTTAAGACAAAATGCAAAAGGAGCTCATCCCATCCGGCAGACCCCTATGTTGGTGTGTTCCAGTATTAC
 AAGATCAACTCTGACCCGTGAAATGTGAGAAGTACACATTCTGTTAAAATACTGCTTTAAGATCTTACCA
 CTCCAATCAATGTTAGTGTGATAGGACATCACCATTCAAGGGGGGGTGTGTTCAACGTCATGGAAGCAGCT
 GACACTCTGAAACTCAGGCAAGGACATGTTGATATTGTTAATTACAATGGAAGTTAAACATTCTTCTGTG
 CACACAAATGGATGGCTCTCTTAAGTGAAGAAAGAGTCAATGAGATTGGCCAGCACATGGAGCTGTAATCCAG
 AGAGAAGGAAAGCTAGAAATTATTAAAAGAATGGACTGTCAGCGAAATCTGTACGGTTCTGTGCAAAGAG
 GTGTTTGCCTGAACTATTTAAGAGACTTTGT

751/237

FIGURE 151

MLNSNVLLWLTALAIKFTLIDSQAQYPVNVNTNYKIRGLRTPNIELGPVEQYLGVPYASP
PTGERRFQPPEPPSSWTGIRNTTQFAAVCPQHLDERSLLHDMLPIWFTANLDTLMTYVQDQN
EDCLYLNIYVPTEDGANTKKNADDITSNDRGEDEDIHDQNSKKPVMVYIHGGSYMGTGNMI
DGSILASYGNVIVITINYRLGILGFLSTGDQAAKGNYGLLDQIQALRWIEENVGAFGGDPKR
VTIFGSGAGASCVSLLTLSHYSEGLFQKAIIQSGTALSSWAVNYQPAKYTRILADKVGCNML
DTTDMVECLRKNKYKELIQQTITPATYHIAFGPVIDGDVIPDDPQILMEQGEFLNYDIMLGV
NQGEGLKFVDGIVDNEDGVTNPNDFSVSNFVDNLYGYPEGKDTLRETIKFMYTDWADKENP
ETRRKTLVALFTDHQWVAPAVAADLHAQYGSPTYFYAFYHHCQSEMKPSSWADSAHDEVYV
FGIPMIGPTELFSCNFSKNDVMLS A VVMTYWTNFAKTGDPNQPVPQDTKFIHTKPNRFEEVA
WSKYNPKDQLYLHIGLKPRVRDHYRATKVAFWLELVPHLN LNEIFQYVSTTKVPPPDMTS
FPYGTRRSPAKIWPTTKRPAITPANNPKHSKDPHKTGPEDTTVLIETKRDYSTELSVTI AVG
ASLLFLNILAFAALYYKKDKRRHETHRRSPQRNTNDIAHIQNEEIMS LQM KQLEHDHECE
SLQAHDTLRLTCPDYTLRRSPDDIPLMTPNTITMIPNTLTGMQPLHTFNTFSGGQNSTN
LPHGHSTTRV

Signal sequence:

amino acids 1-24

Transmembrane domains:

amino acids 189-204, 675-692

152/237

FIGURE 152

GGGAAAGATGGCGGCCACTCTGGGACCCCTGGGTCTGGCAGCAGTGGCGGCATGTTGT
CGGCTCGGGATGGGTCACAGGATGTTACTCCTTCTTGTGTTGGGTCTGGCAGGGCCA
CAGCAAGTCGGGCAGGTCAAACGTTGAGTACTGTAAACAGGGAGCACTCGCTGTCGAAGCC
CTACCAGGGTGTGGCACAGGCAGTTCTCACTGTGAATCTGATGGCAATGCCATGGTGA
TGACCCAGTATATCCGCCTACCCCCAGATATGCAAAGTAACACAGGGTGCCTGTGGAACCGG
GTGCCATGTTCTGAGAGACTGGGAGTTGCAGGGTGCACTCAAAATCCATGGACAAGGAAA
GAAGAATCTGCATGGGATGGCTGGCAATCTGGTACACAAAGGATGGATGCAGCCAGGGC
CTGTGTTGGAAACATGGACAAATTGTGGGGCTGGGAGTATTGTAGACACCTACCCCAAT
GAGGAGAACGAGCAAGAGCGGTATTCCCACATCTCAGGCATGGTAACACGGCTCCCT
CAGCTATGATCATGAGCGGGATGGCGGCCCTACAGAGCTGGGAGGCTGCACAGCCATTGTCC
GCAATCTCATTACGACACCTTCTGGTGAATCGCTACGTCAAGAGGCATTGACGATAATG
ATGGATATTGATGGCAAGCATGAGTGGAGGGACTGCATTGAAGTGCCGGAGTCCGCTGCC
CCCGGGCTACTACTCGGCACCTCCTCCATCACTGGGATCTCTCAGATAATCATGATGTCA
TTTCCTGAAGTGTGTTGAACTGACAGTGGAGAGAACCCAGAAGAGGAAAGCTCCATCGA
GATGTGTTCTGCCCTCAGTGGACAATATGAAGCTGCCTGAGATGACAGCTCCACTGCC
CCTGAGTGGCCTGCCCTTCTCATCGTCTTTCTCCCTGGTGTGTTCTGTATTGCCA
TAGTCATTGGTATCATACTCTACAACAAATGGCAGGAACAGAGCCGAAAGCGCTTCTTACTGA
GCCCTCCTGCTGCCACCACTTGTGACTGTCACCCATGAGGTATGGAAGGAGCAGGCAGTG
GCCTGAGCATGCAGCCTGGAGAGTGTCTGTCTAGCAGCTGGTGGGACTATATTCTG
TCACTGGAGTTGATGCAGGGACCCGCATTCCATGGTGTGCATGGGACATCTA
CTGGTCTGGGAAGCCACCCACCCAGGGCAATGCTGCTGTGATGTGCCTTCCCTGCAGTCC
TTCCATGTGGAGCAGAGGTGTGAAGAGAATTACGTGGTGTGATGCCAAATCACAGAAC
AGAATTTCATAGCCAGGCTGCCGTGTTGACTCAGAAGGCCCTCTACTCAGTTTG
AATCCACAAAGAATTAAAAACTGGTAACACCACAGGCTTCTGACCATTGCTGGGTT
TTGCATTGACCCAAACCCCTGCCTACCTGAGGAGCTTCTTGGAAACCAGGATGGAACACT
TCTCCCTGCCCTACCTCCTTCACTCCATTGCTCTCTGTGTCACCTGAGCTG
GGAAAGGCATTGGATGCCCTCTGTTGGGGCTGGGCTGCAGAACACACCTGCGTTCAC
TGGCCTTCATTAGGTGCCCTAGGGAGATGGCTTCTGCTTGGATCACTGTTCCCTAGCAT
GGGTCTGGGTCTATTGGCATGCCATGCCCTCCAAATCAAGTCTCTCAGGCCCTCAGTG
AAGTTGGCTAAAGGGTGGTGTAAAATCAAGAGAACGCTGGAGACATCATGGATGCCATG
GATTAGCTGCAACTGACCAGCTCCAGGTTGATCAAACAAAAGCAACATTGTGATGTG
GTCTGACCATGTGGAGATGTTCTGGACTTGTCTAGGCCTGCTAGCTGCATGTTGTAGT
TACGATTTGGAATCCACTTTGAGTGTGAAAGTGTAAAGGAAGGCTTCTTCTACACCTT
GGCTTGGATATTGCCAGAGAACATTGGCTTTTTCTTAATGGACAAGAGACAGT
TGCTGTTCTCATGTTCAAGTCTGAGAGCAACAGACCCCTCATCATGTGCCTGGAGAGTT
CACTGTCATTGAGCAGCACGCCAGTGAGTGCCTGCTGTCAACCCATTCTTCACTGCC
TTTGACAAGGGGTTACATGCTGCTCACCTTACTGCCCTGGGATTAAATCAGTTACAGGCCAG
AGTCTCCTGGAGGGCCTGGAACCTGAGTCCTCTATGAACCTCTGTAGCCTAAATGAAAT
TCTAAATCACCAGTGGAAACCAAAAAAAAAAAGGGCGGCCGCAGTCTAGAGTCG
ACCTGCAGTAGGGATAACAGGGTAATAAGCTGGCCGCATGG

753/2 37

FIGURE 153

>/usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50911
><subunit 1 of 1, 348 aa, 1 stop
><MW: 39711, pI: 8.70, NX(S/T): 1
MAATLGPLGSWQQWRRCLSARDGSRMLLLLLLLGSGQGPQQVGAGQTFEYLKREHSLSKPYQ
GVGTGSSSLWNLMGNAMVMTQYIRLTPDMQSKQGALWNRVPCFLRDWELQVHFKIHGQGKKN
LHGDSLAIWYTKDRMQPGPVFGNMDKFVGLGVFVDTYPNEEKQQERVFPYISAMVNNGSLSY
DHERDGRPTELGGCTAIVRNLHYDTFLVIRYVKRHTIMMDIDGKHEWRDCIEVPGVRLPRG
YYFGTSSITGDLSDNHDVISLKLFEVTERTPEEEKLHRDVFLPSVDNMKLPEMTAPLPPPLS
GLALFLIVFFSLVFSVFAIVIGIILYNKWQEQRKRFY

Signal sequence:

amino acids 1-38

Transmembrane domain:

amino acids 310-329

754/237

FIGURE 154

CCGAGCCGGGCGCGCAGCGACGGAGCTGGGCCGCTGGGACCATGGCGTGAGTGCACATACGGATCAGTCT
 CTGATGGTGGGTCTTAACCTCAGTGGGACTCCAAGATTTCATGAAGAAAATCAGTTGCTTCATTCAAGAAT
 TGGGTCTGGCTCAGAATTCTCGCAGCTGGTGAAGAATCTGTTCTAGAAGAGGTTAAATTAATGCCTGAGTCT
 GACATGTTCCGAGTTGAGGTGAAGAGAAAATAGAATACTTAATAATGCTTCCGCAACCGCTTCT
 TGCTGCTGGCCTGGCTGCCTGGCCTTGTGAGCCTCAGCCTGCAGCTTCCACCTGATCCCCTGTGA
 CGACTCTAAGAATGGAATGAGTAGCAAGAGTCAGAAGAGAATCATGCCGACCCCTGTGACGGAGCCCCCTGTGA
 CAGACCCGTTATGAAGCTTGTACTGCAACATCCCAGTGTGGCCAGCGCAGCATGGAAGGTATGCC
 CGCATCATTTAACGCTGGTCTCAGTGCAGTGTCTCGCCACGGAGACAGGTACCCACTGTATGTCATTCCA
 AAACAAAGCAGCACAGAAATTGACTGACTCTGGCTAACAGGAAACCGTATCACCCAAAATGGAAGCTTCA
 TTAGTCACATGTCAAAAGGATCCGGAGCCTTGTGAACTCCCTGCCTTTACCCAAATCACC
 CATGCTGAGATGGGAGAGCTCACACAGACAGGAGTTGTGCAGCATGGTGCAGACGGTCACTGCTGAGGGATA
 TCTATCTAAAGAACACAAACTCCCTGCCAATGATTGGTCTGCAGACCAGCTCATTTAGAGACCACGGGAAAA
 GCCGGACCCCTAACAAAGTGGGCTGGCTTGTCTGGCTTCTCCAGATTGACTGGAAGAAGATTATTC
 GGCACCCAGCCAAGTGGCTGCTGCTGGCTATTGCCCAGTAAGAACCCAGTATCTGAAACCAAACCATCG
 AGCGTCTCAGTACCTCTACGTTGAAAACAGCCAGCTGGAGAAGACCTACGGGAGATGGCAAGATCGTGG
 ATGCCCCACCAAGCAGCTTAGAGCTGCCAACCCCATGACTCCATGCTGCCACTCTGCCACAATGTCAGCT
 TTCCCTGCTACAGAAATGGCTGTGTTGACATGGAGCCTCAAGGTAAATTAGACCCATCAGATCGAGGATGAAA
 GGGAAAGACGGGAGAAGAAATTGACTTCGGGTATTCTCTCTGGGTGCCACCCATCCTGAACCAAACCATCG
 GCCGGATGCGCTGCCACCGAGGGCAGGAAAGAAGAGCTTTGCCCTCTACTCTGCTCATGATGTCACTCTG
 CACCAGTTCTCAGTGCCTGGGCTTCAAGGAGCCAGGTTCCAGGTTGCAGCCAGGGTGTATCTTGTGAGCTT
 GGCAAGACAGAGAAAAGCCCAGTGAACATTCCGCTGGATTCTTACAATGGCGTCATGTCACATTCCACACCT
 CTTCTGCCAAGACACCACAAGCCTCTCCAAGGCCATGTCGGCTGTGAAAACCTGGTCCGCTTGTGAAAA
 GGGACATGTTGTAGCCCTGGGGCAGTGGTACAATTATTATGATGTCATGTCACAGGGAAAGGATTCTAAAGG
 TATGCACTACAGCACTAGAATCCATGCCAATACAGAGCATAGGGAAAGGTCCACTTCTAGTTGCTGTTAC
 TAAGGTTAGAAGATTATTGCTTTAAAGGTAATATTGTTGGAACCACAGATGGTGGGGTTGAACAGT
 AAGCACATTGCTGCAATGGTACGTGAAATTGCTGGTACAAAATGCCAGTTCAACAGAGGAATAGAAGGTACTT
 TATCATAGCCAGACTTCGCTTAGAATGCCAGAATATTAGTTCAAGACCTGAAAGTTGCCAATCCAAGTTGAC
 TCTCTGCCCTGCCCATGTTACTATGTCATGGAACAGCACACCTCAACCAAATTTTTAATCTTAGACATT
 TTTACCTGCTCTGTTAAGAATTCTGAAAGTGTATTATCTAAATAAAGGTTGCCAAACTTTCTGTAAAGG
 GCCAGATTGTAATATTCAAGACTGTGTCAGGACAAAAGGCCACATACAGTCCTGTCATAACTACTCAACTCTG
 TTCTGAAGCAGGAAAGCCACCACAGACAGTACATAAAGGAATATGTCAGCTGGTCTCCAGGCCAGACAAAACA
 GATGGTGACCAAGACTGGCCCTGGGCTGTAGTTGCTGACCCCTCATCTAAAAAATAGGCTATAACTACAAATTG
 ACTTCCAGCACTTTGAGAACGAGGTGAATACAAGAATTATTCAATGGTCCCTCCAGTAACCTCTGCTAGAAACA
 CAGAATTGGTCTGATCTGACACTAGAACAAAATGAGGGTAATTAACATTGAAATTAGAATGAATCATAGAA
 AACTGATTAGAAGAATACCTGATGTTATGATGATTGTTGACAAGAGATGTTAAGTATGTTCTAAATATTG
 CTGCTGTAATGCTATTGCTGTATGCTGAAATTGTTGATGCCATTAGTATTGTTAGTTAGGAAAATATT
 TTCTGAAGCAGCTTGTAGATGACTCTTATTCTGTAGTAATATTCAATTGCTGACCTGCTGGTGGTTAGAAG
 GAGGCTAGAAGATGAATTCAAGGACTTCTTCAAAACTAATTATGGCTCATCCCTTGACAAGCTGAGA
 ACTGGATTCACTTTAACCATTCATCAGTTCAAAATGGTAAATTCTGATTGATTGTTAAATGCGTTTGG
 AGAACATTGCTATTAGGTAGTTACAGATCTTATAAGGTGTTTATATTAGAAGCAATTATAATTACATCTG
 TGATTCTGAACTAATGGTCTAATTCAAGAGAAATGGAAGTGAAGATTCTGTTGTCATGGCATTCC
 AACCTTTCTCTTGTGTTGTCCAGTGTGCAATTGTAATGTCATTCTATAAAATAATTGAAATAA

155/237

FIGURE 155

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48329
><subunit 1 of 1, 480 aa, 1 stop
><MW: 55240, pI: 9.30, NX(S/T): 2
MLFRNRFLLLALAALLAFVSLSLQFFHLIPVSTPKNGSSKSRKRIMPDPVTEPPVTDPVY
EALLYCNIPSVAERSMEGHAPHHFKLVSVHFIRHGDPLYVIPKTKRPEIDCTLVANRKP
YHPKLEAFISHMSKGSGASFESPLNSLPLYPNHPLCEMGELETQTVVQHQLONGQLLRDIYLK
KHKLPLNDWSADQLYLETTGKSRTLQSGLALLYGFLPDFDWKKIYFRHQPSALFCSGSCYCP
VRNQYLEKEQRQQYLLRÍKNSOLEKTYGEMAKIVDVPTKQLRAANPIDSMLCHFCHNVSFPC
TRNGCVDMEHKVIKTHQIEDERERREKKLYFGYSLLGAHPILNQTIGRMQRATEGRKEELF
ALYSAHDVTLSPLSALGLSEARFPRFAARLIFELWQDREKPSEHSVRILYNGVDVTFHTSF
CQDHHKRSPKPMCPLENLVRFVKRDMFVALGGSGTNYYDACHREGF
```

Signal sequence:

amino acids 1-18

156/237

FIGURE 156

AAAAAAGCTCACTAAAGTTCTATTAGAGCGAATACGGTAGATTCCATCCCCTTTGAAGAACAGTACTGTGGA
GCTATTAAAGAGATAAAAACGAAATATCCTTCTGGAGTTCAAGATTGTCAGTAATTGGTAGGACTCTGAGC
GCCGCTGTTCCAATCGGGAGAGAAAAGCGGAGATCTGCTGCCTTGACCGCCTGAAGCACAAACGAGAT
ACCTAGGAATGAACCATCCCTGGGAGTATGTGAAACAACGGAGGAGCTGTGACTTCCAACGTCCCATTCTAT
GGGCGAAGGAACGTCTGACTTCAGTGGTAAGGGCAGAATTGAAAATAATTCTGGAGGAAGATAAGA**ATGAT**
TCCTGCGCAGTGCACCGGGACTACAAAGGGCTTGTCTGCTGGGAATCCTCTGGGACTCTGTGGGAGACCGG
ATGCACCCAGATACTGCTATTCACTGGAGCTGGAGAAAGGCTCTAGGGTGGCGACATCTCCAGGGACCT
GGGGCTGGAGCCCCGGAGCTCGGGAGCGCAGTCCGATCATCCCCAGGGTAGACGGGAGGAGCTGTATGGGGCCATCAAGTG
GAATCCGCGCAGCGGAGCTGGTACCGGGAGGAGATAGACGGGAGGAGCTGTATGGGGCCATCAAGTG
TCAATTAAATCTAGACATTCTGATGGAGATAAGTGAATAATTATGAGTAGAAGTAGAAGTAAGGGACATTAA
CGACAATGCCCTACTTCTGAAAGTGAATTAGAAATAAAAATTAGTGAATAGCAGCCACTGAGATGCGGTT
CCCTCTACCCCAGCCTGGATCCGGATATCGGAAGAACTCTGCAAGAGCTACGAGCTCAGCCGAACACTCA
CTTCTCCCTCATCGTCAAATGGAGCGACGGTAGTAAGTACCCGAATTGGTAGTGAACCGGCCCTGGACCG
CGAAGAAAAGGCTGCTCACCACTGGTCTTACGGCCTCCGACGGGGCAGCCGGTGCACAGGCACCCGCCG
CATCCGCGTGTGGTCTGGATGCCAACGACAACGCCACCGCTTGTCAAGCCGAGTACCCGCCAGGGCTTCC
GGAGAATCTGGCTTGGCAGCGCAGCTGCTTGTAGTCAACGCTACCGCCTGACGAAGGAGTCAATGCGGAAGT
GAGGTATTCTCCGGTATGTGGACGACAAGGGGGGAGTTCAACTAGATGTAAATTCAAGGGACAAATATC
AACAAATAGGGGAGTTGGACGACGAGGAGTCAGGATTCTACCAAGATGGAAGTGAACCAATGGATAATGCAGGATA
TTCTGCGCAGCAAAGTCTGATCACTGTTCTGGACGTGAACGACAATGCCCAAGAGTGGCTCCTACCTCT
CGCCAGCTCGGTTCCGAAAACCTCTCCAGAGGGACATTAATTGCCCTTTAAATGTAATGACCAAGATCTGA
GGAAAACGGACAGGTGATCTGTTCATCCAAGGAATCTGCCCTTTAAATTAGAAAATCTACGGAAATTACTA
TAGTTTAGTCACGACATAGTCTGGATAGGGAACAGGTTCTAGCTACACATCACAGTGACGCCACTGACCG
GGGAACCCGCCCTATCCACGGAAACTCATATCTCGTGAACGTCAGACACCAACGACAACCCGCCGGTCTT
CCCTCAGGCCCTTATCCCGTTATACCCAGAGAACAAATCCAGAGGGAGTTCCCTCGTCTGTGACCCCCA
CGACCCCGACTGTGAAGAGAACGCCAGATCACTTATCCCTGGCTGAGAACACCATCCAAGGGCAAGCCTATC
GTCCTACGTGTCATCAACTCCGACACTGGGTTACTGTATGCGCTGAGCTCCTCGACTACGAGCAGTCCGAGA
CTTGCAAGTGAAGAGTGTGGCGGGACAACGGGACCCGCCCTCAGCAGAACGTGTCGTGAGCCTGTTCGT
GCTGGACCAGAACGACAATGCGCCCGAGATCTGTACCCGCCCTCCGACGGACGGTTCCACTGGCGTGGAGCT
GGCTCCCGCTCCGAGAGCCGGTACCTGGTACCCAAGGTGGCGGTGGACAGAGACTCCGCCAGAACGC
CTGGCTGCTTCTACCGTCTGCTCAAGGCCAGCGAGCCGGACTCTCTCGTGGGGTCTGCACACGGCGAGGTGCG
CACGGCGAGCCCTGTCAGAGACGCGCTCAAGCAGAGCCTGAGTGGCGTCCAGGACACGGCCAGCC
CCCTCTCCGCCACTGTCACGCTCACCGTGGCGTGGCCGACAGCATCCCCAAGTCTGGGGACCTCGGCAG
CCTCGAGTCTCCAGCTAACCTCAGACCTCACTCTGTACCTGGTGTAGCGGTGGCGCGGTCTCCTG
CGTCTTCTGGCTTCGTACCTGCTGCTGGCGCTCAGGCTGCCGCTGGCACAAGTCACGCCCTGCTGCAGGC
TTCAGGAGGGCGCTTGACAGGAGCGCCGGCTCGCAGTTGTGGCGTGGACGGGGTGCAGGGTTCTGTCAGAC
CTATTCCCACGAGGTTTCCCTACCCACGGACTCGCGGAAGAGTCACCTGATCTTCCCCCAGCCAACTATGAGA
CATGCTCGTACGGAGAGCTTGAAAAGCGAGGCCCTTTGCTGTCAGGTGATTCGGTATTTCTAAAGA
CAGTCATGGTTAATTGAGGTGAGTTATATCAAATCTCTTCTTTTTTTAATTGCTGTCTCCAAGC
TGGAGTGCAGCGGTACGATCATAGCTCACTGCCCTCAAACCTCTAGGCTCAAGCAATTATCCACCTTGCCT
CCGGTGTAAACAGGGACTACAGGTGCAAGCCACCTACTGTCTGCCCTATCTATCTATCTATCTATCT
CTATCTATCTATCTATCTATTACTTCTTGTACAAGACGGGAGTCTCACGCCCTGTAATCCAGTACTTGGGAGGC
CGAGGCGGGTGGATCACCTGAGGTGGAGTTGAGACCAGCC**TG**ACCAACATGGAGAAACCCGTCTAACTAA
AAAATACAAAATTAGCGGGCGTGGTGTGATGTCTGTAATCCAGCTACTTGGGAGGCTGAGATTGTGCCATTGCACTCCAGCCTGGCAACAAAGAGTG
AAACTCTATCTCA

1571237

FIGURE 157

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48306
><subunit 1 of 1, 916 aa, 1 stop
><MW: 100204, pI: 4.92, NX(S/T): 4
MIPARLHRDYKGLVLLGILLGTLWETGCTQIRYSVPEELEKGSRVGDISRDLGLEPRELAER
GVRIIPRGRTQLFALNPRSGSLVTAGRIDREELCMGAIKCQILNLDILMEDKVKIYGVEVEVR
DINDNAPYFRESELEIKISENAATEMRFPLPHAWDPDIGKNSLQSYLESPNTHFSLIVQNGA
DGSKYPELVLKRALDREEKAHHVLTASDGGDPVRTGTARIRVMVLDANDNAPAFQAQPEYR
ASVPENLALGTQLLVVNATDPDEGVNAEVRYSFYVDDKAAQVFKLDNSGTISTIGELDHE
ESGFYQMEVQAMDNAGYSARAKVLITVLDVNDNAPEVVLTSLASSVPENS PRGTLIALLNV
DQDSEENGQVICFIQGNLPFKLEKSYGNYYSLVTDIVLDREQVPSYNITVTATDRGTPLLST
ETHISLNVADTNNDNPPVFPQASYSAYIPENNPRGVSLVSVTAHDPCCEENAQITYSLAENTI
QGASLSSYVSINSDTGVLVALSSFDYEQFRDLQVKVMARDNGHPPNSSVSLFVLDQNDN
APEILYPALPTDGSTGVELAPRSAEPGYLVTKVVAVDRDQNAWLQSYRLLKASEPGLFSVG
LHTGEVRTARALLRDALKQSLVVAVQDHGQPPLSATVTVAVADSIPQVLADLGSLES
PA NSETSDLTLYLVAVA AVSCVFLAFVILLALRLRRWHKSRLLQASGGGLTGA
PASHFVGVD
GVQAFLQTYSHEVSLTTDSRKSHLIFPQPNYADMLVSQESFEKSEPLLLSGDSVFSKD
HGLIEVSLYQIFFLFFFNCVSQAGVQRYDHSSLRPQTPLKQLSHLC
RCNRDYRCKPPTVCLS
IYLSIYLSIYLSIYLLSCTDGSLTPVIPVWEAEAGGSPEVGSLRPA

Signal sequence:

amino acids 1-30

Transmembrane domains:

amino acids 693-711, 809-823, 869-888

758/237

FIGURE 158

CCCAAGGCTCTAGTGCAGGAGGAGAAGGAGGAGCAGGAGGTGGAGATTCCCAGTTAAAG
GCTCCAGAACATCGTGTACCAGGCAGAGAACTGAAGTACTGGGGCCTCCTCCACTGGGTCCGAA
TCAGTAGGTGACCCGCCCTGGATTCTGGAAGACCTCACCATGGGACGCCCCGACCTCGT
GCGGCCAAGACGTGGATGTT CCTGCTCTGCTGGGGGAGCCTGGCAGGACACTCCAGGGC
ACAGGAGGACAAGGTGCTGGGGGTATGAGTGCCAACCCATTGCAGCCTGGCAGGC GG
CCTTGTTCCAGGGCCAGCAACTACTCTGTGGCGGTGTCCTTAGGTGGCAACTGGTCCTT
ACAGCTGCCCACTGTAAAAAACGAAATACACAGTACGCCTGGAGACCACAGCCTACAGAA
TAAAGATGCCCAAGAGCAAGAAATACCTGTGGTTCA GTCCATCCCACACCCCTGCTACAACA
GCAGCGATGTGGAGGACCACAACCATGATCTGATGCTTCAACTGCGT GACCAGGCATCC
CTGGGGTCCAAAGTGAAGCCC ATCAGCCTGGCAGATCATTGCACCCAGCCTGGCAGAAGTG
CACCGTCTCAGGCTGGGC ACTGT CACCAGTCCCCGAGAGAATTTCTGACACTCTCAA CT
GTGCAGAAGTAAAATCTTCCCAGAAGAAGTGTGAGGATGCTTACCCGGGCAGATCACA
GATGGCATGGTCTGTGCAGGCAGCAGCAAAGGGCTGACACGTGCCAGGGCATTCTGGAGG
CCCCCTGGTGTGATGGTGC ACTCCAGGGCATCACATCCTGGGCTCAGACCCCTGTGGGA
GGTCCGACAAACCTGGCGTCTATACCAACATCTGCCGTA CCTGGACTGGATCAAGAAGATC
ATAGGCAGCAAGGGCTGATTCTAGGATAAGC ACTAGATCTCCCTTAATAAAACTCACAACTCT
CTGGTTC

759/237

FIGURE 159

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48336
<subunit 1 of 1, 260 aa, 1 stop
<MW: 28048, pI: 7.87, NX(S/T): 1
MGRPRPRAAKTWMFLLLGGAWAGHSRAQEDKVLGGHECQPHSQWPQAALFQGQQLLCGGVL
VGGNWVLTAAHCKPKYTVRLGDHSLQNKGPEQEIPVVQSIPHPCYNSSDVEDHNHDLMLL
QLRDQASLGSKVKPISLADHCTQPGQKCTVSGWGTVTSPRENFPDTLNCAEVKIFPQKKCED
AYPGQITDGMVCAGSSKGADTCQGDSGGPLVCDGALQGITSWGSDPCRSOKPGVYTNICRY
LDWIKKIIGSKG

Important Features:**Signal peptide:**

amino acids 1-23

Transmembrane domain:

amino acids 51-71

N-glycosylation site.

amino acids 110-113

Serine proteases, trypsin family, histidine active site.

amino acids 69-74 and 207-217

Tyrosine kinase phosphorylation site.

amino acids 182-188

Kringle domain proteins motif

amino acids 205-217

760/237

FIGURE 160

GGCGCCGGTGCACCGGGCGGGCTGAGCGCCTCCTGCGGCCGGCCTGCGGCCCGGCCGC
CGCGCCGCCAACGCCCCAACCCCAGGCGCTCCGCCGCGAGGCCCGGCCGCCGCC
GCCCGCGCCAGGTGAGCGCTCCGCCGCGAGGCCCGGCCGCCGCCGCC
CCCCGGCCGGCGGGGAACCGGGCGGATTCTCGCGTCAAACCACTGATCCCATAAAC
ATTCATCCTCCGGCGCCGCGCTGCGAGCGCCCCGCCAGTCGCGCCGCCGCCCTCG
CCCTGTGCGCCCTGCGCGCCCTGCGCACCCCGCGGCCAGGCCAGCCAGAGCGGGCGGAGC
GGAGCGCGCCGAGCCTCGTCCCGCGGCCGGCCGGGCCGTAGCGGCGGCCCTGGA
TGC GGACCCGGCCGCGGGAGACGGGCCCGCCCCGAAACGACTTCA GTCCCCGACGC
CCCGCCAACCCATCGAAGAGGGCGTCCGCTGGAGGGAGCCGGCTGCTGGCATGGTG
CTGTGGCTGCA GGGCTGGCAGGTGGCAGCCCCATGCCAGGTGCCTGCGTATGCTACAATGA
GCCCAAGGTGACGACAAGCTGCCCGAGCAGGGCTGCAGGCTGTGCCGTGGCATCCCTG
CTGCCAGCCAGCGCATCTCCCTGCACGGCAACCGCATCTCGATGTGCCAGCTGCCAGCTC
CGTGCCTGCCGCAACCTCACCATCCTGTGGCTGCACTCGAATGTGCTGGCCGAATTGATGC
GGCTGCCTTCACTGGCCTGGCCCTCTGGAGCAGCTGGACCTCAGCGATAATGCACAGCTCC
GGTCTGTGGACCCTGCCACATCCACGGCCTGGCCGCTACACACGCTGCACCTGGACC
TGC GGCTGCAGGAGCTGGGCCGGGCTGTTCCCGGGCTGGCTGCCCTGCAGTACCTCTA
CCTGCAGGACAACCGCTGCAGGCAGTCGACTGCCTGATGACACCTTCCGCGACCTGGCAACCTCA
CACACCTCTCCTGCACGGCAACCGCATCTCAGCGTGGCCAGCGCGCTTCCGTGGCTG
CACAGCCTCGACCGTCTCCTACTGCACCAAGAACCGCGTGGCCATGTGCACCCGATGCC
CCGTGACCTGGCCGCTCATGACACTCTATCTGTTGCCAACAACTATCAGCGCTGCC
CTGAGGCCCTGGCCCCCTGCGTGCCCTGCAGTACCTGAGGCTCAACGACAACCCCTGGTG
TGTGACTGCCGGCACGCCACTCTGGCCTGGCTGCAGAAGTCCGGCTCCTCCTCGA
GGTCCCTGCAGCCTCCGCAACGCCCTGGCTGGCGTACCTCAAACGCCCTAGCTGCC
ACCTGCAGGGCTGCCTGTGGCACCGGCCCTAACATCCATCTGGACCGCAGGGCACC
GATGAGGAGCCGCTGGGCTTCCAAGTGCTGCCAGCCAGATGCCGCTGACAAGGCC
ACTGGAGCCTGGAAGACCAGCTCGGCAGGCAATGCGCTGAAGGGACCGTGGCC
ACAGCCCAGGGCAACGGCTCTGGCCACGGCACATCAATGACTCACCC
CCTGGCTCTGCTGAGCCCCGCTCACTGCAGTGCAGGCCAGGGCTCCGAGC
CCCCACCTCGGGCCCTGCCGGAGGCCAGGCTGTTCACGCAAGAACGC
GCCGTCTGGGCCAGGCAGGCAAGCAGGGACTGGTACTCAGAAGGCT
CTACCCAGCCTCACCTGCAGCCTCACCCCCCTGGCCTGGCGTGGT
TGGGCCCTGCTGACCCCCAGCGGACACAAGAGCGTCTCAGCAGCCAGGT
GGGTCTCTCCACGCCGCAAGCCAGCCGGGGCGACCC
GTCCTCCCTGATGGACGCCTGCCGCCACCC
TTCGGCGGCAGCGTTGTTCCAGAACGCCCTCCACCC
GCATTTATTTACTGTGAAAAATACGGACGACGT
AAAA

7611237

FIGURE 161

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44184
><subunit 1 of 1, 473 aa, 1 stop
><MW: 50708, pI: 9.28, NX(S/T): 6
MKRASAGGSRLLAWVLWLQAWQVAAPCPGACVCYNEPKVTTSCPQQGLQAVPVGIPAASQRI
FLHGNRISHVPAASFRCRNLTILWLHSNVLARIDAAAFTGLALLEQLDLSDNAQLRSVDPA
TFHGLGRLHTLHLDRCGLQELGPGLFRGLAALQYLYLQDNALQALPDDTFRDLGNLTHLFLH
GNRISSVPERAFRGLHSLDRLLLHQNRVAHVPHAFRDLGRLMTLYLFANNLSALPTEALAP
LRALQYLRLNDNPWVCDCRARPLWAWLQKFRGSSSEVPCSLPQRLAGRDLKRLAANDLQGCA
VATGPYHPIWTGRATDEEPLGLPKCCQPDAAVKASVLEPGRPASAGNALKGRVPPGDSPGN
GSGPRHINDSPFGTLPGSAEPPLTAVRPEGSEPPGFPSTGPRRRPGCSRKNRTRSHCRLGQA
GSGGGGTGDSEGSGALPSLTCSTPLGLALVLWTVLGPC
```

Important features:**Signal peptide:**

amino acids 1-26

Leucine zipper pattern.

amino acids 135-156

Glycosaminoglycan attachment site.

amino acids 436-439

N-glycosylation site.

amino acids 82-85, 179-183, 237-240, 372-375 and 423-426

VWFC domain

amino acids 411-425

162/237

FIGURE 162

GGAAGTCCACGGGAGCTGGATGCCAAAGGGAGGACGGCTGGTCCTCTGGAGAGGACTAC
TCACTGGCATATTCTGAGGTATCTGTAGAATAACCACAGCCTCAGATACTGGGACTTAC
AGTCCCACAGAACCGTCTCCAGGAAGCTGAATCCAGCAAGAACAATGGAGGCCAGCGGA
AGCTCATTGCAGACAAGGCAAGTCCTTTTCCCTTCTCCTTTGGCTATCTCTGGCG
GGCGCGCGGAACCTAGAAGCTATTCTGTGGTGGAGGAAACTGAGGGCAGCTCCTTGTCAC
CAATTAGCAAAGGACCTGGGTCTGGAGCAGAGGAAATTCTCAGGGCAGCTCCTTGTCAC
TTCCAGAGGAAACAAACTACATTGCAGCTCAATCAGGAGACCGCGGATTGTTGCTAAAT
GAGAAATTGGACCGTGAGGATCTGTGCGGTACACAGAGCCCTGTGTGCTACGTTCCAAGT
GTGCTAGAGAGTCCTTCGAGTTTCAAGCTGAGCTGCAAGTAATAGACATAACGACC
ACTCTCCAGTATTCTGGACAAACAAATGTTGGTGAAGTATCAGAGAGCAGTCCTCTGGG
ACTACGTTCCCTCTGAAGAATGCCAAGACTTAGATGTAGGCCAAACAAATATTGAGAACTA
TATAATCAGCCCCAACTCCTATTTCGGGTCTCACCGCAAACGCAGTGTGAGCAGGAAAT
ACCCAGAGCTGGTGTGGACAAAGCCTGGACCGAGAGGAAGAAGCTGAGCTCAGGTTAAC
CTCACAGCACTGGATGGTGGCTCTCCGCCAGATCTGCAGTGTACATCGAAGT
CCTGGATGTCAACGATAATGCCCTGAATTGAGCAGCCTTCTATAGAGTGCAGATCTCTG
AGGACAGTCCGGTAGGCTCTGGTGTGAAGGCTCTGCCACGGATGTAGACACAGGAGTC
AACGGAGAGATTCCATTCACTTTCCAAGCTCAGAAGAGATTGGAAAACCTTAAGAT
CAATCCCTGACAGGAGAAATTGAACAAAAAAACAACCTGATTCGAAAACCTCAGTCCT
ATGAAGTCAATATTGAGGCAAGAGATGCTGGAACCTTTCTGGAAAATGCACCGTTCTGATT
CAAGTGTAGATGTGAACGACCATGCCAGAAGTTACCATGTCATGCCATTACAGCCAAAT
ACCTGAGAACGCGCTGAAACTGTGGTGCACCTTCAGTGTTCAGATCTGATTAGGAG
AAAATGGGAAAATTAGTTGCTCATTCAAGGAGGATCTACCCCTCCTGAAATCCGCGGAA
AACTTTACACCCACTAACGGAGAGACCACAGACAGAGAAAGCAGAGCGGAATACAACAT
CACTATCACTGCACTGACTTGGGGACCCCTATGCTGATAAACACAGCTCAATATGACCGTGC
TGATGCCGATGTCAATGACAACGCTCCGCCCTCACCCAAACCTCTACACCCCTGTC
CGCGAGAACACAGCCCGCCCTGCACATCCGCAGCGTCAGCGCTACAGACAGAGACTCAGG
CACCAACGCCAGGTCAACCTACTCGCTGCTGCCGCCAGGACCCGCACCTGCCCTCACAT
CCCTGGTCTCCATCAACGCCAACCGCCACCTGTTGCCCTCAGGTCTCTGGACTACGAG
GCCCTGCAGGGTTCCAGTCCCGCTGGCGCTTCAGACCAACGGCTCCCCGGCGCTGAGCAG
CGAGGCCTGGTGCCTGGTGTGGACGCCAACGACAACCTGCCCTCGTGTGCTTAC
CGCTGCAGAACGGCTCCCGCCCTGCACCGAGCTGGTGCCCCGGCGAGCCGGCTAC
CTGGTGACCAAGGTGGTGGCGTGGACGGCAGCTGGGCCAGAACGCCCTGGCTGTCGTACCA
GCTGCTCAAGGCCACGGAGCTGGTCTGTTCCGGCTGTGGCGCACAAATGGCGAGGTGCGCA
CCGCCAGGCTGCTGAGCGAGCGCACGCCAACGCTGACGTGCTCTGGTGACGGCTTCTC
AATGGCGAGCCTCCCGCCTGCCACGCCAACGCTGACGTGCTCTGGTGACGGCTTCTC
CCAGCCCTACCTGCCCTCTCCGGAGGCCGCCAGCCAGGCCAACGCCGACTTGCTCACCG
TCTACCTGGTGGTGGCTGGCTCGGTGTCTCGCTCTTCCCTTTGGTGTGCTCTGGTCTTC
GTGGCGGTGCGGCTGTGTAGGAGGAGCAGGGCAGGCTCGGTGGTGTGCTTGGTCTGG
GGGCCCCCTCCAGGGCATCTGTGGACATGAGCGGCACCCAGGACCTATCCCAGAGCTACC
AGTATGAGGTGTGTGCTGGCAGGAGGCTCAGGGACCAATGAGTCAAGTTCTGAAGCCGATT
ATCCCCAACTCCCTCCCCAGTGCCCTGGAAAGAAATACAAGGAAATTCTACCTTCCAA
TAACCTGGGTTCAATATTCAAGTGACCCATAGTTGACTTTACATTCCATAGGTATT
TGTGGCATTCCATGCCAATGTTATTCCCCCAATTGTTGTAATATTGTACGGAT
TTACTCTTGATTCTCATGTTCTCCCTTGTAAAGTGAACACATTACCTTATT
CCTGGTTCTT

7631237

FIGURE 163

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48314
<subunit 1 of 1, 798 aa, 1 stop
<MW: 87552, pI: 4.84, NX(S/T): 5
MEASGKLICRQRQVLFSLLLGLSLAGAAEPRSYSVVEETEGSSFVTNLAKDLGLEQREFSR
RGVRVVSRGNKLHLQLNQETADLLLNEKLDREDLCGHTEPCVLRFQVLLESPFEFFQAELOV
IDINDHSPVFLDKQMLVKVSESSPPGTTFPLKNAEDLDVGQNNIENYIISPNSYFRVLTRKR
SDGRKYPELVLDKALDREEEAELRLTLTALDGGSPPRSFTAQVYIEVLDVNDAPEFEQPFY
RVQISEDSPVGFLVVVKVSATDVDTGVNGEISYSLFQASEEIGKTFKINPLTGEIELKKQLDF
EKLQSYEVNIEARDAGTFSGKCTVLIQVIDVNDHAPEVTMSAFTSIPENAPETVVALFSVS
DLDSENGKISCSIQEDLPFLLKSAENFYTLTERPLDRESRAEYNITITVTDLGTPMLITQ
LNMTVLIADVNDNAPAFQTQSYTLFVRENNSPALHIRSVSATDRDGTNAQVTYSLLPPQDP
HLPLTSLVSINADNGHLFALRSLDYEALQGFQFRVGASDHGSPALSSEALVRVVVLDANDNS
PFVLYPLQNGSAPCTELVPRAAEPGYLVTKVVAVDGDSGQNAWLSYQLLKATELGLFGVWAH
NGEVRTARLLSERDAAKHRLVVLVKDNCEPPRSATATLHVLLVDGFSQPYLPLPEAAPTQAQ
ADLLTVYLVVALASVSSLFLFSVLLFVAVRLCRRSRAASVGRCLVPEGPLPGHLVDMSGTRT
LSQSYQYEVCLAGGSGTNEFKFLKPIIPNFPPQCPGKEIQGNSTFPNNFGFNIQ
```

Important features:**Signal peptide:**

amino acids 1-26

Transmembrane domain:

amino acids 685-712

Cadherins extracellular repeated domain signature.

amino acids 122-132, 231-241, 336-346, 439-449 and 549-559

ATP/GTP-binding site motif A (P-loop).

amino acids 285-292

N-glycosylation site.

amino acids 418-421, 436-439, 567-570 and 786-789

169237

FIGURE 164

ACCCACGCGTCCGCCAACGCGTCCGCCAACGCGTCCGCCAACGCGTCCGCCGTAGCCGTGC
GCCGATTGCCTCTCGGCCTGGCAATGGTCCGGCTGCCGGTCGACGACCGCCCCGCGTCAT
GCGGCTCCTCGGCTGGTGGCAAGTATTGCTGTGGTGCTGGACTTCCGTCCGGCGTGG
AGGTTGCAGAGGAAGTGGTCGCTTATGGTCAGAGGAGCAGCCTGCTCACCTCTCCAGGTG
GGGGCTGTGTACCTGGGTGAGGAGGAGCTCCTGCATGACCCGATGGCCAGGACAGGGCAGC
AGAAGAGGCCAATGCGGTGCTGGACACCCAAGGCGATCACATGGTGTGCTGCTG
TGATTCCCTGGGAAGCTGAGGACAAAGTGAGTTCAGAGCCTAGCGCGTCACCTGTGGTGC
GGAGGAGCGGAGGACTCAAGGTGCAACGTCCGAGAGAGCCTTCTCTGGATGGCGCTGG
AGCACACTCCCTGACAGAGAAGAGGAGTATTACACAGAGCCAGAAGTGGCGGAATCTGACG
CAGCCCCGACAGAGGACTCCAATAACACTGAAAGTCTGAAATCCCCAAAGGTGAACGTGAG
GAGAGAAACATTACAGGATTAGAAAATTCACTCTGAAAATTAAATATGTCACAGGACCT
TATGGATTCTGAACCAAACGGTAGTGACTGTACTCTAGTCCTGTTACACCCCGTGGT
GCCGCTTCTGCCAGTTGGCCCTCACTTTAACCTCTGCCCCGGCATTCCAGCTCTT
CACTTTGGCACTGGATGCATCTCAGCACAGCAGCCTTCTACCAGGTTGGCACCGTAGC
TGTTCTTAATATTTATTATTCAAGGAGCTAAACCAATGGCCAGATTAAATCATAACAGATC
GAACACTGGAAACACTGAAAATCTCATTAAATCAGACAGGTATAGAAGCCAAGAAGAAT
GTGGTGGTAACTCAAGCCGACCAAATAGGCCCTTCCCAGCACTTGATAAAAAGTGTGGA
CTGGTTGCTGTATTTCTTATTCTTTAATTAGTTTATTATGTATGCTACCATTGAA
CTGAGAGTATTCGGTGGCTAATTCCAGGACAAGAGCAGGAACATGTGGAGTAGTGATGGTCT
GAAAGAAGTTGGAAAGAGGAACCTCAATCCTCGTTCAAGAAATTAGTGCTACAGTTCTATA
CATTTCTCCAGTGACGTGTTGACTTGAAACTCAGGCAGATTAAAAGAATCATTGTTGAA
CAACTGAATGTATAAAAAAAATTATAAACTGGTGTAACTAGTATTGCAATAAGCAAATGC
AAAAATATTCAATAG

165/237

FIGURE 165

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48333
><subunit 1 of 1, 360 aa, 1 stop
><MW: 39885, pI: 4.79, NX(S/T): 7
MVPAAGRRPPRVMRLLGWWQVLLWVLGLPVRGVEVAEESGRLWSEEQPAHPLQVGAVYLGE
ELLHDPMGQDRAAEEANAVGLDTQGDHMVMLSVPGEAEDKVSSEPSGVTCGAGGAEDSRC
NVRESLFLSLDGAGAHFPDREEEYYTEPEVAESDAAPTEDSNNTESLKSPKVNCERNITGLE
NFTLKILNMSQDLMDFLNPNQSDCTLVLFYTPWCRCFSASLAPHFNSLPRAFPALHFLALDAS
QHSSLSTRFGTVAVPNILLFQGAKPMARFNHTDRTLETLKIFIFNQTGIEAKKNVVVTQADQ
IGPLPSTLIKSVDWLLVFSLLFISFIMYATIRTESIRWLIPGQEQUEHVE
```

Important features:

Signal peptide:

amino acids 1-25

Transmembrane domain:

amino acids 321-340

Homologous region to dilsufide isomerase

amino acids 212-302

N-glycosylation site.

amino acids 165-168, 181-184, 187-190, 194-197, 206-209, 278-281
and 293-296

Thioredoxin domain

amino acids 211-227

1661237

FIGURE 166

CCCGGCTCCGCTCCCTCTGCCCTCGGGGTGGGGCCCCACGATGCTGCAGGGCCCTGGCT
CGCTGCTGCTCTCCTCGCCTCGCACTGCTGCCCTGGCTCGGCCGCGGGCTCTCCTC
TTTGGCCAGCCCAGCTTCTCCTACAAGCGCAGCAATTGCAAGCCCATCCCGTCAACCTGCA
GCTGTGCCACGGCATCGAATACCAAGAACATGCCAACCTGCTGGCCACGAGACCA
TGAAGGAGGTGCTGGAGCAGGCCGGCGCTTGGATCCCGCTGGTCATGAAGCAGTGCCACCCG
GACACCAAGAAAGTCTGTGCTCGCTCTCGCCCCGTCTGCCCTGATGACCTAGACGAGAC
CATCCAGCCATGCCACTCGCTCGGTGCAGGTGAAGGACCGCTGCCGGTCATGTCG
CCTCGGCTCCCCGGCCGACATGCTTGAGTGCACCGAGGAAGCTCAAAGGTATG
TGCATCCCCCTCGCTAGCAGCACCCCTGCCAGCCACCGAGGAAGCTCAAAGGTATG
TGAAGCCTGCAAAATAAAATGATGATGACAACGACATAATGAAACGCTTGTAAGGAT
ATTTGCACTGAAAATAAAAGTGAAGGAGATAACCTACATCAACCGAGATAACCAAATCATC
CTGGAGACCAAGAGCAAGACCATTACAAGCTGAACGGTGTGTCGAAAGGGACCTGAAGAA
ATCGGTGCTGGCTCAAAGACAGCTTGCAGTGACCTGTGAGGAGATGAACGACATCAACG
CGCCCTATCTGGTCATGGACAGAAACAGGGTGGGAGCTGGTATCACCTCGTGAAGCGG
TGGCAGAAGGGCAGAGAGAGTTCAAGCGCATCTCCCGACAGCAGCTGACCGTGAAGCTTA
GTCCCGGCATCCTGATGGCTCCGACAGGCCGCTCCAGAGCACGGCTGACCGTGAAGCT
GGGATCTCAGCTCCGTTCCCAAGCACACTCCTAGCTGCTCCAGTCTCAGCCTGGCAGCT
TCCCCCTGCCCTTGACGTTGCATCCCCAGCATTCTGAGTTATAAGGCCACAGGAGTG
GATAGCTGTTTACCTAAAGGAAAAGCCCACCCGAATCTGTAGAAATATTCAAACATA
AAATCATGAATATTTAA

1671237

FIGURE 167

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50920
><subunit 1 of 1, 295 aa, 1 stop
><MW: 33518, pI: 7.74, NX(S/T): 0
MLQGPGSLLLLFLASHCCLGSARGLFLFGQPDSYKRSNCKPIPVNLQLCHGIEYQNMRLPN
LLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLCSLFAPVCLDDLDETIQPCHSLCVQVKDR
CAPVMSAFGFPWPDMLECDRFPQDNDLCIPLASSDHLLPATEEAPKVCEACKNKNDDNDIM
ETLCKNDFALKIKVKEITYINRDTKIILETKSHTIYKLNGVSERDLKKSVLWLKDSDLQCTCE
EMNDINAPYLVMGQKQGGELVITSVKRWQKGQREFKRISRSIRKLQC

Important features:

Signal peptide:

amino acids 1-20

Cysteine rich domain, homologous to frizzled N terminus
amino acids 6-153

768/237

FIGURE 168

GTGGAGGCCGCGACGATGGCGGGCCGACGGAGGCCGAGACGGGTTGGCCGAGCCCCGGG
CCCTGTGCGCGCAGCGGGCCACCGCACCTACGCGCCGCTGGGTGTTCTGCTCGCGATC
AGCCTGCTCAACTGCTCCAACGCCACGCTGTGGCTCAGCTTGACCTGTGGCTGACGTCAT
TGCTGAGGACTTGGTCCTGTCCATGGAGCAGATCAACTGGCTGTCACTGGTCTACCTCGTGG
TATCCACCCCATTGGCGTGGCGGCATCTGGATCCTGGACTCCGTGGGCTCCGTGCGGCG
ACCATCCTGGGTGCGTGGCTGAACTTGCCGGAGTGTGCTACGCATGGTGCCCTGCATGGT
TGTTGGGACCCAAAACCCATTGCCTTCCTCATGGGTGGCCAGAGCCTCTGTGCCCTGCC
AGAGCCTGGTCATCTCTCCAGCCAAGCTGGCTGCCCTGTGGTCCCAGAGCACCAAGCGA
GCCACGGCCAACATGCTGCCACCATGTCGAACCCCTGGCGTCCTGTGGCCAATGTGCT
GTCCCCTGTGCTGGTCAAGAAGGGTGAGGACATTCCGTTAATGCTCGGTGTCTATACCATCC
CTGCTGGCGTCGTCTGCCCTGTCCACCATCTGCCCTGTGGGAGAGTGTGCCCTGCC
CCCTCTGCCGGGCTGCCAGCTCACCTCAGAGAAAGTCCCTGGATGGCTCAAGCTGCAGCT
CATGTGGAACAAGGCCTATGTCATCCTGGCTGTGCTTGGGGAAATGATCGGGATCTCTG
CCAGCTTCTCAGCCCTCCTGGAGCAGATCCTCTGTGCAAGCGGCCACTCCAGTGGTTTCC
GGCCTCTGTGGCGCTCTTCATCACGTTGGATCCTGGGGCACTGGCTCTGGCCCTTA
TGTGGACGGACCAAGCACTCACTGAGGCCACCAAGATTGGCCTGTGCCTGTTCTCTGG
CCTGCGTGCCCTTGCCCTGGTGTCCCAGCTGCAGGGACAGACCCCTGCCCTGGCTGCCACC
TGCTCGCTGCTCGGCTGTTGGCTTCGGTGGGCCATGGAGTTGGCGTGA
GTGTTCTTCCCGTGGGGAGGGGCTGCCACAGGCATGATCTTGTGCTGGGAGGCC
AGGGAATACTCATGCTGGCAATGACGGCACTGACTGTGCGACGCTCGAGCCGCTT
TCCACCTGCCAGCAGGGGAGGATCCACTGACTGGACAGTGTCTGCTGCTGATGCCGG
CCTGTGCACCTTCTCAGCTGCATCCTGGCGGTCTTCCACACCCATACCGCGCCTGC
AGGCCGAGTCTGGGGAGCCCCCTCCACCGTAACGCCGTGGCGGCCAGACTCAGGCCG
GGTGTGGACCGAGGGGAGCAGGAAGGGCTGGGTCTGGGGCCAGCACGGGACTCCGGA
GTGCACGGCGAGGGGGCCTCGTAGAGGACCCAGAGGGCCGGAGCCCCACCCAGCCT
GCCACCGAGCGACTCCCCGTGCGCAAGGCCAGCAGCCACCGACGCCCTCCGCC
AGACTCGCAGGCAGGGTCCAAGCGTCCAGGTTATTGACCCGGCTGGGTCTCACTCCT
CTCCTCCCCGTGGGTGATCACGTAGCTGAGCGCCTGTAGTCCAGGTTGCCGCCACATCGA
TGGAGGCGAACTGGAACATCTGGTCCACCTGCAGGGCGGGGGCGAAAGGGCTCCTGCGGGCT
CCGGGAGCGAATTACAAGCGCGCACCTGAAAA

769/237

FIGURE 169

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50988
><subunit 1 of 1, 560 aa, 1 stop
><MW: 58427, pI: 6.86, NX(S/T): 2
MAGPTEAETGLAEPRALCAQRGHRTYARRWFLLAISLLNCSNATLWLSFAPVADVIAEDLV
LSMEQINWLSLVYLVVSTPFGVAIWIIDSVGLRAATILGAWLN FAGS VL R M V PCMV VGTQN
PFAFLMGGQSLCALAQSLVIFSPAKLAALWFPEHQRATANMLATMSNPLGV LV ANV LSPVLV
KKGEDIPLMLGVYTIPAGVVCLLSTICLWESVPPTPPSAGAASSTSEKFLDGLKLQLMWNK
YVILA VCLGGMIGISASFSALLEQILCASGHSSGFSGLCAL FITFG I GL GAL ALGP YV DRTK
HFTEATKIGLCLFSIACVPFALVSQLQQT LALAATCSLLGLFGF SVGPVAME LAVECSFPV
GE GAATGM IFV LG QAE G ILIM LAM TALT VRR SE PSL STC QQ GEDPLD WTV SLLL MAGL CTFF
SCILA VFF HTPY RRL QAES GE PPSTRNA VGGADSGPGVDRGGAGRAGV LG PSTAT PECTARG
ASLEDPRGP GSPHPACHRATPRAQPAATDAPS R P GRL AGRV QASRF IDPAGSHSSFSSPWVIT

Important features:

Signal peptide:

amino acids 1-44

Transmembrane domains:

amino acids 61-79, 98-112, 126-146, 169-182, 201-215, 248-268,
280-300, 318-337, 341-357, 375-387, 420-441

N-glycosylation site.

amino acids 40-43 and 43-46

Glycosaminoglycan attachment site.

amino acids 468-471

770/237

FIGURE 170

GTCCCCACATCCTGCTCAACTGGGTCAAGGTCCCTTAGACCAGCTTGTCCATCTTGCTGAAGTGGACCAAC
 TAGTTCCCAGTAGGGGGCTCCCTGGCAATTCTTGATCGCGTTGGACATCTCAGATCGCTCCAATGAAGA
 TGGCCTTGCCTGGGTCTCTGCTTCTATAATCATCTAACTATGGGACAAGGTTGTGCCGGCAGCTCTGGGG
 AAGGAGCACGGGCTGATCAAGCCATCCAGGAAACACTGGAGGACTTGTCAGCCTGAAAGAAACTCTAGTGGT
 TCTGAATCTAGCCCACCTGGCGTAAGC**ATGAT**GCAACTCTGCAACTCTGCTGGGGCTTGGGGCCAGGTGG
 CTACTGGTACAGTGTGGGAGGTGACCACTCAGGTGAAATACCAAGTGTCAAGGAAAGTGGC
 ATCTGGTACAGTGTGGGAGGTGACCACTCAGGTGAAATACCAAGTGTCAAGGAAAGTGGC
 CCAGGTGTTGCAGCTGCCCTCAGCGCTCCCCATTAGGTGAACTCTGAGGAGGTGCTCAGCACAGGCAGGCG
 GCTGGATCGAGAGCAGCTGTGCCGACAGTGGGATCCCTGCCCTGGTTCTTGTGATGTGCTGCCACAGGGGATT
 GGCTCTGATCCATGTGGAGATCCAAGTGCTGGACATCAATGACCACCCAGGTTCCCAAAGGCGAGCAGGA
 GCTGGAAATCTTGAGAGCGCTCTCGGAACCCGATCCCCCTGGACAGAGCTTGTGACCCAGACACAGGCC
 TAACACCTGACACACCTACACTCTGCTCTCCAGTGAGCACTTGTGCTTGGATGTGATTGTGGGCCCTGATGAGAC
 CAAACATCGAGAACACTCATAGTGGTAAGGAGCTGGACAGGAAATCCATCTCATTTGTGATCTGGTGTAACTGC
 CTATGACAATGGGAAACCCCCCAAGTCAGGTACAGCTGGTCAAGGTCAACGCTTGGACTCCAATGACAATAG
 CCCTCGTTGTGAGAGTCACTGGCACTGGAAATCCAAGAAGATGCTGACCTGTTGACTACGCTTCTCATAAAAT
 GACCGCCACAGACCCGACCAAGGCCCCATGGGGAGGTGGAGTTCTCTCTGAGTAAGCACATGCCCTCAGAGGT
 GCTGGACACCTCAGTATTGATGCCAAGACAGGCCAGGTCTCTGCGTCACCTCTAGACTATGAAAAGAACCC
 TGCCCTACGAGGTGGATGTTCAAGCAAGGGACCTGGGCTCCAAATCTATCCAGCCATTGCAAAGTTCTCATCAA
 GGTCTGGATGTCATGACAACATCCAAGCATTCCACGTACATGGGCCCTCCAGGCATCACTGGTGTCAAGGC
 TCTTCCAAAGGACAGTTTATTGCTCTGTGATGGCAGATGACTTGGGATCTAGGACACAATGGTTGGTCAACTG
 CTGGCTGAGCCAAGAGCTGGGCCACTCAGGTGAAAAGAACTAATGCCAACACATGATGTTGCTAACCAATG
 CACACTGGACAGAGCAGTGGCCAAATATACCCCTACTCTGTTAGGCCAAGACCAAGGACTCAGGCCCTATC
 AGCCAAGAAACAGCTAGCATTCAAGTCAACTGACATCAACGACAATGCACTGTGTTGAGAAAAGCAGGTATGA
 AGTCTCACGGGGAAAACAACATTACCTCTCTTACCTCATTACCATCAAGGCTCATGATGCACTTGGCATT
 TAATGGAAAAGTCTCATACCGCATCCAGGACTCCCCAGTTGTCACCTAGTAGCTATTGACTCCAACACAGGAGA
 GGTCACTGTCAGAGGTCACTGAACATGAAGAGATGGCCGCTTGGAGTTCCAGGTGATGCAAGGACAGCGG
 GCAACCCATGCTGCATCCAGTGTCTGTGTTGGCTCAGGCCCTCTGGATGCCAATGATAATGCCCAAGGGTGGT
 CCAGGCTGTGCTCAGCGATGGAAAAGGCCAGGCCCTCTGGTGTGAATGCCCTCACAGGCCACTGCTGGTGC
 CATCGAGACTCCAATGGCTTGGGCCAGCGGGCACTGACACACCTCAGTGGCACTCACAGCTCCGGCATT
 CCTTTGACAACCATGTGGCAAGAGATGCAACTCGGGGCAATGGAGAGGCCCTCTACAGCATCCGAATGG
 AAATGAAGGCCACCTCTCATCCTCACCCCTCATCGGGGCACTGTTGTCATGTCACCAATGCCAGCAGCCT
 CATTGGGAGTGAGTGGAGACTGGAGATAGTAGTAGAGGACCAGGGAAAGCCCCCCTACAGACCCGAGGCCCTGTT
 GAGGGTCACTGTTGTCACCAAGTGTGGACCACTGAGGGACTCAGCCGCAAGCCTGGGCCCTGGAGCATGTCGAT
 GCTGACGGTGTCTGGCTGTACTGTTGGCATCTCGGGTGTGATCTGGCTTGTGTCATGTCACATCTGCC
 GACAGAAAAGGAGACAACAGGGCTCAACTGTCGGGGAGGGCAGTCCACCTACGGCCAGCAGGCCAAGAGGCC
 CCAGAACACATTCAAGGCCAGACATCCACCTCTGCTGCTCAGGGTCAAGGAGGTGAGCCTGTGAAGT
 CGGGCAGTCCCACAAAGATGTGGACAAGGAGGGCATGATGGAAGCAGGCTGGACCCCTGCCCTGCAAGGCC
 CCACCTCACCCGACCCGTACAGGACGCTGCGTAATCAAGCAACCAGGGAGCACCGCGGAGAGCCGAGAGGT
 GCTGCAAGACACGGTCAACCTCTTCAACCATCCCAGGCAGAGGAATGCCCTCCGGAGAACCTGAAACCTTCC
 CGAGGCCAGGCCACAGGCCAGTTCAGGCCTCTGAAAGGTTGCAAGGCAGGCCACAGGGAGGCTGGC
 TGGAGACCAAGGGCAGTGAGGAAGGACCTCTGTGAGAAGCACTGCTAGAAGAAGAGCTGTCAAGTCTGCTGGACCC
 CAGCACAGGTGTGGCCCTGGACCGGCTGAGGCCCTGACGGCCCTGGATGGCGAGACTCTCTTGGCC
 CACCAACTACCGTACAATGTGATCTCCCGGATGCTGCAAGGCCAGGGAGGGAGGACTTCCAGACGTTG
 CAAGGCAGAGGCCACAGGCTGAGGCCAACAGGCAGGCCAGGGCTGGCCAGCACCTTGTCTGGAGATGAGCTCACT
 GCTGGAGATGCTGCTGAAACAGCAGCTCCAGCATGCCGTGGAGGCCCTGGAGGCCCTGCGGGCTCTCGG
 CTGGGGAGGACCCCTAGTTAGACTTGGCACCACTGCAAGCAGGCTCAGGCATGAAAGTGCAGGGACCCAGGTGG
 AAAGACGGGACTGAGGGCAAGAGCAGAGGCAAGCAGCAGCAGCAGCAGGCTGCTG**TGA**ACATACCTCAGACGCC
 CTGGATCCAAGAACCAAGGGCCTGAGGATCTGTGACAAGAGCTGGTTCTAAATCTTGTAACACTAGCTAG
 CGGCCGCTGAGAACCTTAAAGGGTACTGATGCTACCCCCACAGAGGAGGCAAGAGGCCAGGACTAACAGCTGAC
 TGACCAAAGCAGCCCCCTGTGAGCAGCTGTGAGCTTGGAGGAGCAGGGACGGTTGTGGCTGAGATAAGTGT
 TCCTGGCAAAACATATGTGGAGCACAAGGGTCAAGTCTCTGGCAGAACAGATGCCACGGAGTATCACAGGCAG
 AAAGGGTGGCTTCTGGTAGCAGGAGTCAGGGGCTGTACCTGGGGTGCAGGAAATGCTCTGACCTAT
 CAATAAGGAAAAGCAGTAAAAA

771/237

FIGURE 171

```

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48331
<subunit 1 of 1, 1184 aa, 1 stop
<MW: 129022, pI: 5.20, NX(S/T): 5
MMQLLQLLLGPGGGYLFLGDCQEVTTLTVKYQVSEEVPSGTIVGKLSQELGREERRQA
GAAFQVLQLPQALPIQVDSEEGLLSTGRRLDREQLCRQWDPCLVSFDVLATGDLALIHVEIQ
VLDINDHQPRFPKGEOQELEISESESALRTRIPLDRALDPDTGPNTLHTYTLSPSEHFAVDVIV
GPDETAKHAELIVVKELDREIHSFFDLVLTAYDNGNPPKSGTSLVKVNVLDSNDNSPAFAESS
LALEIQEDAAPGTLLIKLTATDPDQGPNGEVEFFLSKHMPPEVLDTSIDAKTGQVILRRPL
DYEKNPAYEVDVQARDLGPNPIPAHCKVLIKVLVDVNDNIPSIHVTVASQPSLVSEALPKDSF
IALVMADDLDSGHNGLVHCWLSQELGHFRKRTNGNTYMLLTNATLDREQWPKYTLTLLAQD
QGLQPLSAKKQLSIQISDINDNAPVFEKSRYEVSTRENNLPSIHLITIKAHADLGINGKVS
YRIQDSPVAHLVAIDSNTGEVTAQRSLNYEEMAGFEFQVIAEDSGQPMLASSVSVWVSSLDA
NDNAPEVVQPVLSDGKASLSVLVNASTGHLLVPIETPNGLGPAGTDTPPLATHSSRPFLTT
IVARDADSGANGEPLYSIRNGNEAHLFILNPHTGQLFVNVTNASSLIGSEWELEIVVEDQGS
PPLQTRALLRVMFVTSDHLRDSARKPGALSMSMLTVICLAVLLGIFGLLALFMSICRTEK
KDNRAYNCREAESTYRQQPKRPQKHIQKADIHLVPVLRGQAGEPCEVGQSHKDVDKEAMMEA
GWDPCLOQAFHLPFLYRTLNRNQGNQGAPAESREVQDVTNLLFNHPRQRNASRENLNLPF
QPATGQPRSRLKVAGSPTGRLAGDQGSEEAPQRPPASSATLRRQRHNGKVSPKEESGPRQ
ILRSLVRLSVAFAERNPVEELTVDSPPVQQISQLLSQLHQGQFQPKPNHRGNKYLAKPGGS
RSAIPDTDGPSARAGGQTDPEQEEGPLDPEEDLSVKQLLEEELSSLLDPSTGLALDRLSAPD
PAWMARLSSLPLTTNYRDNVISPDAATEEPRTFQTFGKAEAPELSPTGTRLASTFVSEMSSL
LEMILLEQRSSMPVEAASEALRRLSVCGRTLSLDLATSAASGMKVQGDPGGKTGTEGKSRGSS
SSSRCL

```

Important features:**Signal peptide:**

amino acids 1-13

Transmembrane domain:

amino acids 719-739

N-glycosylation site.

amino acids 415-418, 582-585, 659-662, 662-665 and 857-860

Cadherins extracellular repeated domain signature.

amino acids 123-133, 232-242, 340-350, 448-458 and 553-563

172/237

FIGURE 172

CGGACGCGTGGCGGACGCGTGGGGAGAGCCGCAGTCGGCTGCAGCACCTGGAGAAGG
CAGACCGTGTGAGGGGCCTGTGGCCCAGCGTGCTGTGGCTCGGGAGTGGAAAGTGGAG
GCAGGAGCCTCCTAACACTTCGCCATGAGTTCCATCGACTCCAGCATCATGATTACCT
CCCAGATACTATTTGGATTGGCTTTCTTCATCGCCAATTGTTAAAGACTAT
GAGATACGTCACTATGTTGACAGGTGATCTCTCCGTGACGTTGCATTTCTGCACC
GTTGAGCTCATCATCTTGAAATCTTAGGAGTATTGAATAGCAGCTCCGTTATTTCACT
GGAAAATGAACCTGTGTGAATTCTGCTGATCCTGGTTCATGGTGCCTTTACATTGGC
TATTTATTGTGAGCAATATCCGACTACTGCATAAACAAACGACTGCTTTCCCTGTCT
ATGGCTGACCTTATGTATTCTGGAAACTAGGAGATCCCTTCCATTCTCAGCCAA
AACATGGGATCTTATCCATAGAACAGCTCATCAGCCGGTTGGTGTGATTGGAGTGACTCTC
ATGGCTCTTCTTCTGGATTGGTGTCAACTGCCATACACTACATGTCTTAC
CAGGAATGTGACTGACACGGATTCTAGCCCTGGAACGGCGACTGCTGCAAACCATGGATA
TGATCATAAGCAAAAGAAAAGGATGGCAATGGCACGGAGAACATGTTCCAGAAGGGGAA
GTGCATAACAAACCATCAGGTTCTGGGAATGATAAAAGTGTACCACCTCAGCATT
AAGTAAAATCTTACTCTTATTCAACAGGAAGTGGATGCTTGGAAAGAATTAAGCAGGCAGC
TTTTCTGGAAACAGCTGATCTATGCTACCAAGGAGAGAACATGAGTCTTGGAA
AAGGGAAATTTAATTCTGGTACTTTCTCTATTTACTGTGTTGGAA
CATGGCTACCACATATTGTTGATCGAGTTGGAAAACGGATCCTGTACAAGAGGCA
TTGAGATCACTGTGAATTATCTGGAAATCCAATTGATGTGAAGTTGGTCCAACACATT
TCCTCATTCTGTTGAAATAATCATCGTCACATCCATCAGAGGATTGCTGATCACTCT
CAAGTTCTTATGCCATCTCTAGCAGTAAGCCTCAATGTCATTGCTCTGCTATTAGCAC
AGATAATGGCATGTACTTGTCTCCTCTGTGCTGATCCGAATGAGTATGCCATTAGAA
TACCGCACCATAAATCACTGAAGTCCTGGAGAACTGCAGTTCAACTCTATCACCGTTGGT
TGATGTGATCTCCTGGTCAGCGCTCTCTAGCATACTCTCTATTGGCTACAAAC
AGGCACCAAGAGAACATGGCACTGAACTTAAGCCTACTACAGACTGTTAGAGGCCAGT
GGTTCAAAATTAGATATAAGAGGGGGAAAATGGAACCAGGGCTGACATTTATAAAC
AAACAAAATGCTATGGTAGCATTTCACCTCATAGCATACTCCTCCCCGTAGGTGATA
CTATGACCACATGAGTAGCATGCCAGAACATGAGAGGGAGAACACTCAAGACAATACTCA
GCAGAGAGCATCCGTGTGGATATGAGGCTGGTAGAGGGGGAGAGGCCAGAACACTAA
AGGTAAAAATACACTGGAACTCTGGGAAGACATGTCTATGGTAGCTGAGCCAAACACGT
AGGATTTCCGTTAAGGTTACATGGAAAAGGTTATAGCCTTGCTGAGATTGACTCATT
AAAATCAGAGACTGTAACAAAAAAAAAAAAAGGGCGGCCGACTCTAGAGTCG
ACCTGCAGAAGCTGGCCGCATGGCCAACCTGTTATTGCAGCTTATAATG

773/237

FIGURE 173

MSFLIDSSIMITSQILFFFGFWLFFMRQLFKDYEIRQYVVQVIFSVDFAFSCTMFELIIFEI
LGVLNSSSRYFHWKMNLCVILLILVFMVPFYIGYFIVSNIRLLHKQRLLFSCLLWLTMYFF
WKLGDPFPILSPKHGILSIEQLISRGVIGVTLMALLSGFGAVNCPTYMSYFLRNVTDI
LALERRLLQTMDMIISKKKRMAMARRTMFQKGEVHNKPSPGFWMIKSVTTSAASSENLTLIQ
QEVDALLELSRQLFLETADLYATKERIEYSKTFKGKYFNFLGYFFSIYCVWKIFMATINIVE
DRVGKTDPVTRGIEITVNYLGIQFDVKFWSQHISFILVGIIIVTSIRGLLITLTKFFYAISS
SKSSNVIVLLAQMIMGMYFVSSVLLIRMSMPLYRTIITEVLGELQFNFYHRWFDVIFLVSA
LSSILFLYLAHKQAPEKQMAP

Important features:**Signal peptide:**

amino acids 1-23

Potential transmembrane domains:amino acids 37-55, 81-102, 150-168, 288-311, 338-356, 375-398,
425-444**N-glycosylation sites.**

amino acids 67-70, 180-183 and 243-246

Eukaryotic cobalamin-binding proteins

amino acids 151-160

7741237

FIGURE 174

CATGGGAAGTGGAGCCGGAGCCTCCTTACACTGCCATGAGTTCTCATCGACTCCAGCA
TCATGATTACCTCCNGANACTATTTTGAGTTGGGTGGCTTCNGGCCAATGTT
TAAAGACTATGAGATACTGTCAGTATGTTGACNGGTGATCTCTCCGTGACGTTGCCATTT
CTTGCACCATGTTGAGCTCATCTTGAAATCTNGGAGTATTGAATAGCAGCTCCGT
TATTTCACTGGAAAATGAACCTGTGTAAATTCTGCTGATCCTGGTTNTCATGGTGCCTTT
TTACATTGGCTATTTTATTGTGAGCAATATCCGACTACTGCATAAACAAACGACTGCTTTTT
CCTGTCTCTTATGGCTGACCTTATGTATTTCCAG

775/237

FIGURE 175

GTGTTGCCCTGGGGAGGGAAAGGGAGCCNGGCCTTCCTAAATTTGCCAAGGGTTTC
TTTNTGAATTCCGGGTTNNGNATACCTCCCAGAAAATATTTTGAGTTGGGTAGNTT
TTTTCATGCGCCAATTGTTAAAGACTATGAGATACTGAGCTATGTTGACAGGTGATNTT
NTCCGTGACGTTGCATTTCTGCACCAGTTGAGCTCATCATNTTGAAATNTTAGGAG
TATTGAATAGCAGCTCCCGTTATTCACGGAAAATGAACTGTGTAATTCTGCTGATC
CTGGTTTCATGGTGCCTTTACATTGGCTATTTATTGTGAGCAATATCCGACTACTGCA
TAAACAACGACTGCTTTCTGTCTNTTATGGCTGACCTTATGTATTNTNTGGAAAN
TAGGAGATCCCTTCCCATTCTC

7961237

FIGURE 176

CTCGGCCAGGGATCGTCCC**ATGGCCGGGCTCGGAGCCGCACCCCTGGGGGGCTCCGGGATTGCTACCTTT**
 TGGCTCCCTGCTCGTCAACTGCTCTCTCACGGCTGTGCCCTCAATCTGGACGTGATGGGTGCCTGCCAA
 GGAGGGCGAGCCAGGCAGCCTCTCGGCTCTGTGGCCCTGCACCCGGCAGTTGCAAGCCCCAACCCAGAGCTG
 GCTGCTGGTGGTGCCTCCCAGGCCCTGGCTCTCCTGGCAGCAGGCAATCGCACTGGAGGCCCTTCGCTTG
 CCCGTTGAGCCTGGAGGAGACTGACTGCTACAGAGTGGACATCGACCAGGGAGCTGATATGAAAAGGAAAGCAA
 GGAGAACCAGTGGTGGAGTCAGTGTCCGGAGCCAGGGCCCTGGGGCAAGATTGTTACCTGTGCACACCGATA
 TGAGGCAAGGCAGCGAGTGGACCAGATCTGGAGACGCAGGATATGATGGTCGCTGTTGTGCTCAGCCAGGA
 CCTGGCCATCCGGATGAGTGGATGGTGGGGAAATGGAAGTTCTGTGAGGGACGCCCAAGGCCATGAACAATT
 TGGGTCTGCCAGCAGGGCACAGCTGCCCTCTCCCTGATGCCACTACCTCTCTTGGGGCCCAAGGAAC
 CTATAATTGAGGGCACGGCCAGGGTGGAGCTGTGACAGGGCTCAGGGCACACCTGGACAGCG
 TCCCTACGAGCCGGGGGAGAGAAGGAGCAGGACAGGCCCTCATCCGGCTCCTGCCAACAGCTACTTGGCTT
 CTCTATTGACTCGGGGAAAGGTCTGGTGCCTGCAGAAGAGCTGAGCTTGTGAGGCCCTGGCCACAGCTACTTGGCTT
 CAAGGGTGCTGTGGTACATCCTGCGCAAGGACAGCGCCAGTCGCCTGGTGGCCAGGGTATGCTGCTGGGGAGCG
 CCTGACCTCCGGCTTGCTACTCACTGGCTGTGGCTGACCTCAACAGTGTGGCTGGCCAGACCTGATAGTGGG
 TGGGGGGCTACTCTTGAGGCCAACAGAGCTGGGGGTGCTGTGATGTGACTTGAAACCAAGGGGGTCACTG
 GGCCTGGGATCTCCCTCTCCGGCTCTGCGCTCCCTGACTCCATGTTGGGATCAGCCTGGCTGCTGGGGGA
 CCTCAACCAAGATGGCTTCCAGATATTGCACTGGTGGCCCTTGTGATGGTATGGGAAAGTCTCATCTACCA
 TGGGAGCAGCTGGGGGTGTCGCAAACCTCACAGGTGCTGGAGGGCAGGGCTGGGGCATCAAGAGCTCGG
 CTACTCCCTGTCAGGCAGCTTGGATATGGATGGAACCAATACCTGACCTGCTGGGGCTCCCTGGCTGACAC
 CGCAGTGTCTTCAGGGCCAGACCCATCCTCCATGTCTCCATGAGGTCTCTATTGCTCCACGAAGCATCGACCT
 GGAGCAGCCAACTGTGCTGGCGGCACTCGGCTCTGTGAGCCTAAGGGTCTGTTCAAGCTACATTGCACTGCCC
 CAGCAGCTATAGCCCTACTGTGGCCCTGGACTATGTTAGATGCGGACACAGACCCGGAGGCTCCGGGGCAGGT
 TCCCCGTGTGACGTTCTGAGCGTAACTGGAGAACAAACCTGACCCAGGACACCAGGGCTCGGGCTGAGCA
 CCAGCATGACCGAGTCTGGAGACGCCATGTTCCAGTCCACGGCTGAGGAAATGTCAAAAGAACGCTCGGGCATTGT
 AGTGACCTTGCTTACAGTCTCCAGACCCCTGCTCCGGGACAGGCTCTGGCCAGGGCTGCCCTCAGTGGC
 CCCCATCCTCAATGCCAACAGCCAGCAGGGCAGAGATCCACTTCCCTGAAGCAAGGCTGTGGTGAAGA
 CAAGATCTGCAGAGCAATTCGAGCTGGTCCACGCCGCTCTGTACCCGGGTGAGCAGCACCGAACCTCAACC
 TCTGCCCATGGATGGAGAACACAGCCCTGTTGCACTGAGTGGCAGCCAGTCATTGGCTGGAGCTGAT
 GGTCAACCTCCAGTCTGACGAGCCGTCTTCATTGAGCTGCCACTGTCCATTGAGGAATGGCATTCC
 CCAGCAACTCTCTCTGGTGTGGAGGGGGAGAGGCCATGCACTGAGCTGAGCCGGATGTGGGAGCAAGGT
 CAAGTATGAGGTCAAGGTTTCAACCAAGGCCAGTCGCTCAAGACCCCTGGCTCTGCCCTCCTCAACATCATGTG
 GCCTCATGAGATGCCAAAGGGAGTGGTCTGCTACCTGGATGAGCTGGAGGGGGGGGGGGGGGGGGGGGG
 GCAGAAAGGGTTGCTCTCCAGGCCAACATCTCCACCTGGATGAGCTGGAGAGCTAGGGATAGGAGGGGGGG
 GCTGGAGCCACCTGAGCAGCAGGGCTGGTGAAGGGCAGGGAGCCAGCATGCTGGTGGCCAGTGTCTCTG
 TGAGAAGAAGAAAACATACCCCTGGACTGCGCCGGGGCACGGCCAAGTGTGTTGAGCTGCCACTCTA
 CAGCTTGTGAGCGCGGGCTGTGCTGCATGTCGTTGGGCCGTCTGGAACAGCACCTTCTGGAGGAGTACTCAGC
 TGTGAAGTCCCTGGAAGTGGTATGGTCCGGCCAACATCACAGTGAAGTCTCCATAAAGAACATTGATGCTCCGAGA
 TGCCTCCACAGTGTGATGGTATACTGGGACCCCATGGCTGTGGTGGCAGAAGGGAGTGGCCCTGGGG
 CATCTCTGCTGTACTGGCTGGCTGCTGGTGTAGCAGCTGCTGGTGGCTCTGGAGATGGGGATTCTT
 CAAACGGCGAAGCACCCGAGGCCACCGTGCCAGGGCAGGGCAGGCCAGGGCAGGCCACTCT
 GTCAAGGAGGAGAAGACGGGACCCATCCTGAGGAACAACAGGGCAGGCCAGGGCAGGCCACTCT
 CCCCCATCCTGGCTGCTGACGGGCACTCCGAGCTGGGCCCCGATGGGACATCCAGGGCAGGCCACTCT
 CATGTCCTGAGCTGGCTGTGGCTGCCCTCCATCCCTCCAGAGATGGCTCTGGGATGAAAGGGTAGAGT
 GGGCTGCTGGTGTGCAAGATTGGCAGGGCTCTCAGGGGACAGACCTCCACCCACAAGAAC
 TCCTCCCACCAACTTCCCTTAGAGTGTGAGATGAGAGTGGGAAATCAGGGACAGGGCCATGGGGATGGGG
 TGAGAAGGGCAGGGGTGCTCTGATGCAAAGGTGGGGAGAAGGGATCTTAATCCCTCCATTACCCCTGT
 GTAACAGGACCCCAAGGACCTGCCCTCCCGGAAGTGCCTTAACCTAGAGGGTGGGGAGGAGGTGTCAGTGA
 CTCAGGCTGCTCTCTAGTTCCCTCTCATCTGACCTAGTTGCTGCCATCAGTCTAGTGGTTCTG
 TTCGTCTATTAAAAAATTTGAGAACAAAAAAAAAAAAAA

771237

FIGURE 177

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA55737
><subunit 1 of 1, 1141 aa, 1 stop
><MW: 124671, pI: 5.82, NX(S/T): 5
MAGARSRDPWGASGICYLFGSLLVELLFSRAVAFNLDVMGALRKEGEPGSLFGFSVALHRQL
QPRPQSLLVGAPQALALPGQQANRTGGLFACPLSLEETDCYRVDIDQGADMQKESKENQWL
GVSVRSQGGPGKIVTCAHRYEARQRVDQILETRDMIGRCFVLSQDLAIRDELDGGEWKFC
RPOGHEQFGFCQQGTAAAFSPDSHYLLFGAPGTYNWKGATARVELCAOGSADLAHLDGPYEA
GGEKEQDPRLIPVPANSYFGFSIDSGKGLVRAEELSFVAGAPRANHKGAVVILRKDSASRLV
PEVMLSGERLTSGFGYSLAVADLNSDGWPDLIVGAPYFFERQEELGGAVVYLNQGGHWAGI
SPLRLCGSPDSMFGISLAVLGDLNQDGFPDIAVGAPFDGDKVFIYHGSSLGVVAKPSQVLE
GEAVGIKSFGYSLSGSLDMGNQYPDILVGSЛАDTAVLFRARPILHVSHEVSIAPRSIDLEQ
PNCAGGHHSVCDLRCFSYIAVPSSYPTVALDYVLDADTDRLRGQVPRVTFLSRNLEEPK
HQASGTVWLKHQHDRVCGDAMFQLQENVKDKLRAIVTLYSLQTPLRRQAPGQGLPPVAP
ILNAHQPSTQRRAEIHFQKQGCEDKICQSNLQLVHARFCTRSDTEFQPLPMVDGTTALFA
LSGQPVIGLELMVNLPSDPAQPQADGDDAHEAQLLVMLPDSLHYSGVRAIDPAEKPLCLSN
ENASHVECELGNPMKRGQAQVTFLYLSTSGISIETTELEVELLLATISEQELHPVSARARVF
IELPLSIAGMAIPQQLFFSGVVRGERAMQSERDVGSKVKEYTVSNQGQSLRTLGS AFLNIM
WPHEIANGKWLlyPMQVELEGGQGPGQKGLCSPRPNILHLDVDSRDRRRELEPPPEQQEPGE
RQEPMMSWVPVSSAEKKKNITIDCARGTANCVVFCPLYSFDRAAVLHVWGRLWNSTFLEEY
SAVKSLEVIVRANITVKSSIKNLMLRDASTVIPVMVYLDPMAVVAEGVPWWVILLAVLAGLL
VLALLVLLWKMGFKRAKHPEATVPQYHAVKIPREDRQQFKEEKTGTILRNNWGSPRREGP
DAHPILAADGHPELGPDPGHPGPFTA

Important features:**Signal peptide:**

amino acids 1-33

Transmembrane domain:

amino acids 1040-1062

N-glycosylation sites.

amino acids 86-89, 746-749, 949-952, 985-988 and 1005-1008

Integrins alpha chain proteins.amino acids 1064-1071, 384-408, 1041-1071, 317-346, 443-465, 385-
407, 215-224, 634-647, 85-99, 322-346, 470-479, 442-466, 379-408
and 1031-1047

7781 237

FIGURE 178

CGCGCCGGGCGCAGGGAGCTGAGTGGACGGCTCGAGACGGCGCGGTGCAGCAGCTCCAGA
AAGCAGCGAGTTGGCAGAGCAGGGCTGCATTCCAGCAGGAGCTGCAGCACAGTGCTGGCT
CACAAACAAGATGCTCAAGGTGTCAGCCGTACTGTGTGTGCAGCCGTTGGTGCAGTCAGTCA
GTCCTCGCAGCTGCCGCGGGTGGCTGCAGCCGGGGCGGTGGACGGCGGTAAATTTC
TGGATGATAAACAAATGGCTACCCACAATCTCTCAGTATGACAAGGAAGTCGGACAGTGGAAC
AAATTCCGAGACGAAGTAGAGGGATGATTATTCCGCACTGGAGTCCAGGAAAACCCTCGA
TCAGGCTTAGATCCAGCTAAGGATCCATGCTAAAGATGAAATGTAGTCGCCATAAAGTAT
GCATTGCTCAAGATTCTCAGACTGCAGTCTGCATTAGTCACCGGAGGCTTACACACAGGGATG
AAAGAACAGGAGTAGACCATAGGCAGTGGAGGGTCCCATTATCACCTGCAAGCAGTG
CCAGTGGCTATCCCAGCCCTGTTGTGGTCAGATGGTCATACCTACTCTTTCAGTGCA
AACTAGAATATCAGGCATGTGCTTAGGAAAACAGATCTCAGTCAAATGTGAAGGACATTGC
CCATGTCCTCAGATAAGCCCACCAAGTACAAGCAGAAATGTTAAGAGAGCATGCAGTGACCT
GGAGTTCAGGGAAGTGGCAAACAGATTGCGGGACTGGTTCAAGGCCCTCATGAAAGTGGAA
GTCAAAACAAGAACAGACAAAAACATTGCTGAGGCCCTGAGAGAACAGATTGATACCACATC
TTGCCAATTGCAAGGACTCACTGGCTGGATGTTAACAGACTTGATACAAACTATGACCT
GCTATTGGACCAAGTCAGAGCTCAGAACAGCTTACCTTGATAAGAACAGTGTACCAAGG
CATTCTCAATTCTGTGACACATACAAGGACAGTTAACAGTCAATAATGAGTGGTGTAC
TGCTTCCAGAGACAGCAAGACCCACCTGCCAGACTGAGCTCAGCAATATTCAAGCGGCA
AGGGTAAAGAACAGCTCCTAGGACAGTATATCCCCCTGTGTGATGAAGATGGTTACTACAAGC
CAACACAATGTCATGGCAGTGGACAGTGCTGGTGTGACAGATATGAAATGAAGTC
ATGGGATCCAGAATAAATGGTGTGCAAGATTGTGCTATAGATTTGAGATCTCCGGAGATTT
TGCTAGTGGCGATTTCATGAATGGACTGATGATGAGGATGATGAAGACGATATTGAATG
ATGAAGATGAAATTGAAGATGATGATGAAGATGAAGGGATGATGATGATGGTGGTGTGAC
CATGATGTATAACATTTGATTGATGACAGTTGAAATCAATAAATTCTACATTCTAATATT
CAAAATGATAGCCTATTAAAATTATCTTCTCCCCAATAACAAAATGATTCTAAACCTCA
CATATATTGTATAATTATTGAAAAATTGCAGCTAAAGTTAGAACATTGTTAAC
AAGAACATTTGCTTGAGTTTATATTCTTACACAAAAAGAAAATACATATGCAGTCTA
GTCAGACAAAATAAAGTTTGAAGTGCTACTATAATAAATTTCACGAGAACAAACTTGT
AAATCTTCCATAAGCAAAATGACAGCTAGTGCTGGATCGTACATGTTAATTGGAAAG
ATAATTCTAAGTGAATTAAAATAAAATTGTTAATGACCTGGGTCTAAGGATTAGG
AAAAATATGCATGCTTAATTGCATTCCAAAGTAGCATCTGCTAGACCTAGATGAGTCAG
GATAACAGAGAGATAACCACATGACTCCAAAAAAAAAAAAAA

779/237

FIGURE 179

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49829
><subunit 1 of 1, 436 aa, 1 stop
><MW: 49429, pI: 4.80, NX(S/T): 0
MLKVS A V L C V C A A W C S Q S L A A A A V A A A G G R S D G G N F L D D K Q W L T T I S Q Y D K E V G Q W N K F R
D E V E D D Y F R T W S P G K P F D Q A L D P A K D P C L K M K C S R H K V C I A Q D S Q T A V C I S H R R L T H R M K E A
G V D H R Q W R G P I L S T C K Q C P V V Y P S P V C G S D G H T Y S F Q C K L E Y Q A C V L G K Q I S V K C E G H C P C P
S D K P T S T S R N V K R A C S D L E F R E V A N R L R D W F K A L H E S G S Q N K K T K T L L R P E R S R F D T S I L P I
C K D S L G W M F N R L D T N Y D L L L D Q S E L R S I Y L D K N E Q C T K A F F N S C D T Y K D S L I S N N E W C Y C F Q
R Q Q D P P C Q T E L S N I Q K R Q G V K K L L G Q Y I P L C D E D G Y Y K P T Q C H G S V G Q C W C V D R Y G N E V M G S
R I N G V A D C A I D F E I S G D F A S G D F H E W T D D E D D I M N D E D E I E D D D E D E G D D D G G D D H D V Y I

Important features:**Signal peptide:**

amino acids 1-16

Leucine zipper pattern.

amino acids 246-267

N-myristoylation sites.

amino acids 357-362, 371-376 and 376-381

Thyroglobulin type-1 repeat proteins

amino acids 353-365 and 339-352

7801237

FIGURE 180

CAGACTCCAGATTCCCTGTCAACCACGAGGAGTCCAGAGAGGAAACGGCGAGCGGAGACAACAGTACCTGACGC
 CTCTTCAGCCGGGATGCCCGCAGCAGGGATGGCGACAAGATCTGGCTGCCCTTCCCCTGCTCCTTCTGGCC
 GCTCTGCCTCCGGTGCCTGCCTGGGCGGCCGCTCACACCTTCCCCTGATAGGCAGTTACCTTACCTTACCCCTT
 CCCGCCGCCAGAAGGAGTGCCTACCCAGCCATGCCCTGAAGGCCCTGCGTGGAGATCGAGTACCAAGTCTTA
 GATGGAGCAGGATTAGATATTGATTTCATCTGGCTCTCCAGAAGGAAAACCTTAGTTTGACAATAACATCAGCACCATT
 TCAGATGGAGTTCACACTGTAGAGACTGAAGTTGGTGAATTACATGTTCTGCTTGAACAATAACATCAGCACCATT
 TCTGAGAAGGTGATTTCTTGAATTAACTGGATAATATGGAGAACAGGACAAGAACAGAAGATTGGAAG
 AAATATATTACTGGCACAGATATATTGGATATGAAACTGGAAGACATCTGGAATCCATCAACAGCATTCAAGTCC
 AGACTAAGCAAAGTGGGCACATACAAATTCTGCTTAGAGCATTGAGCTCGTATCGAAACATACAAGAAC
 AACTTTGATAGAGTCATTCTGGTCTATGGTTAATTAGTGGTATGGTGGTGTGCAGCCATTCAAGTCTTA
 ATGCTGAAGAGTCTGTTGAAGATAAGAGGAAAAGTAGAACATTAAACTCCAAACTAGAGTACGTAACATTGAAA
 AATGAGGCAATAAAATGCAATAACTGTTACAGTCAGAACATTAATGGCTTCTCCAAAATATTGAGATATA
 AAAGTAGGAAACAGGTATAATTAACTGTAAGGAAATTAGTCTTCACTGTGCAAGTAATCTGCTGATCCAG
 TTGACTTAAGTGTGTAACAGGAATATTGCAAGAATATAGGTTAAGTGAATGAAGCCATATTAAACTGCAT
 TTCTCTAATTGAAAAATTGCAAAATGTCTTAGGTGATTAAATAATGAGTATTGGGCTAATTGCAACACC
 AGTCTGTTTAACAGGTTCTATTACCCAGAACCTTTGTAATGCCAGTACAAATTAACTGTGGAAGTTT
 TCAGTTTAAGTTATAAATCACCTGAGAATTACCTAATGATGGATTGAAATAAATCTTAGACTACAAAGCCAA
 CTTTTCTCTATTACATATGCATCTCTCTATAATGTAATAGAATAATAGCTTGAATTAACATTAGGTTTG
 AGATTTTATAACCAAATACATTTCAGTGTAAACATTAGCAGAACAGCTTAGTCTTGTACTTTGCTTACATT
 CCAAAGCTGACATTTCAGGTTCTAAAAACACAAAGTTACACTTACTAAATTAGGACATGTTCTCTTTG
 AAATGAGAATATAGTTAAAGCTTCTCCATAGGGACACATTCTCTAACCCCTAACTAAAGTGTAGGA
 TTTAAAATTAAATGTGAGGTTAAAGTTATTGTAATAGTATCTGCAAGTTAATATCTGCAACAGTTAA
 TAATCATGTTATGTTAATTAAACATGATTGCTGACTTGGATAATTCAATTACAGCAGTTATGAGGAAATA
 TTGCTAAATGATCTGGGCTTACCAATAAATATCTCCTTCTGAGCTCTAAGAATTATCAGAAAACAGGAA
 AGAATTAGAAAAACTTGAGAAAACCTAATCCAATAAAATCACTTAAGTGAAGACTATAAAATATCTAGA
 ATCTGACTGGCTCATCATGACATCCTACTCATAACATAAACTCAAGGAGATGATTAAATTCCAGTTAGCTGGAAG
 AAACCTTGGCTTAGGTTTATTCTACAAAGAATTCTGGTGAATTATTGTAAGCAGGTACATTGTTATA
 AAATGTAAGCCTACTGTAAAGTTAGCAGTTGGTGTACATATTATTAAATTATTAAACACTTTTAT
 TAAAATGGCCTTCTGAACACTTATTGATGTTGAAGTAAGGATTAGAAACATAGACTCCAAGTTTAA
 CACCTAAATGTGAATAACCCATATACAAACAAAGTTCTGCCATCTAGCTTTGAAAGTCTATGGGGCTTAC
 TCAAGTACTAGTAATTAACTTCATCATGAATGAACTATAATTGTAAGTTATGCCATTATAACGTTGTTAT
 GACTACATTGAGTTAGAAACAACTTAAAGTTGGGTATAGAACCCCTAACAGGTTAGTAATGCTGGAATT
 CTTGATGAGCAATAATGATAACCAAGAGACTGATTCTACACTCATAGTAACTGAGTATAAAAGAGATACATTCC
 TCTAGGCCCCCTGGAGAACAGCAGCTTAGATTCTACTGGCAAGGTTTAAAATGAGGTTAAATGCCGTAT
 ATGATCAATTACCTTAATTGCCAAGAAAATGCTTCAGGTGCTAGGGGTATCCTCTGCAACACTTGAGAACAA
 AGGTCAATAAGATCCTGCCTATGAATACCCCTCCCTTGGCTGTTAAATTGCAATGAGAACAAATTACA
 GTACCATAACTATAAAAGCAGGGTACAGATATAAAACTACTGCATCTTCTACCTAACATGAAATTGTTACATAATCTCT
 CCTCTCCTGTATGGCTGTTACTGTACTCTGACTCTTACCTAACATGAAATTGTTACATAATCTCT
 ACATGTATGATTGTCGCACTGATCTAAACCTATGATTCACTTACCATATAAAACGATAATTGCTT
 TATTGAAAAGAATTAGGAATACTAAGGACAATTATTGTTATAGACAAAGTAAAAGACAGATATTAAAGAGG
 CATAACAAAAAAAGCAAAATGTAACAGAGTAAAATCTTAATATTCTAAAGACATACTGTTATCTGCTT
 CATAATGCTTTTAAATTCACTATTCCATTCTAAATTAAAGTTATGCTAAATTGAGTAAGTGTGTTACACTT
 AACAGCTATTGCTTTCAATATACAAATTAAACTACAAATTGTTACTACAATATTAAACTAAGGCCAACGATTTC
 CATAATGTAGCAGTTACCGTGTACCTCACACTAACGGCTAGAGTTGCTCTGATATGCATTGGATGATTAA
 GTTATGCTGTTCTTCATGTGAATGTCAGAACATGGAGGGTTGTTGAAATTGTTATGTAATTAAATCCTTCTTA
 CACATAATGGTGTCTAAAATTGACAAAAAAATGAGCACTTACAAATTGTTATGCTCCTCAAATGAGATTCTTAT
 GTGAAATTAAAGACATTGATTCGCATGTAAGGATTGTTATCTGAAGTACAATAATGCAACATCAGTGTG
 CTCAAACCTGCTTATACATTAAACAGCCATCTTAAATAAGCAACGTATTGAGTAAGTATGATATGATAATAAA
 AAAATTACAAAGGAAA

7/1/237

FIGURE 181

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52196
><subunit 1 of 1, 229 aa, 1 stop
><MW: 26017, pI: 4.73, NX(S/T): 0
MGDKIWLPFPVLLAALPPVLLPGAAAGFTPSLSDFTFTLPAGQKECFYQPMPPLKASLEIEY
QVLDGAGLDIDFHLASPEGKTLVFEQRKSDGVHTVETEVGDYMFCFDNTFSTISEKVIFFEL
ILDNMGEQAQEQQEDWKKYITGTDILDMLKLEDILESINSIKSRLSKSGHIQILLRAFEARDRN
IQESNFDRVNFWSMVNLVVVVSAIQVYMLKSLFEDKRKSRT

Important features:**Signal peptide:**

amino acids 1-23

Transmembrane domain:

amino acids 195-217

N-myristoylation site.

amino acids 43-48

Tyrosine kinase phosphorylation site.

amino acids 55-62

782 / 237

FIGURE 182

CCATCCCTGAGATCTTTATAAAAAACCCAGTCTTGCTGACCAGACAAAGCATAACCAGAT
CTCACCAAGAGAGTCGCAGACACT**ATG**CTGCCTCCCAGGGCCCTGCCAGTGTGTCCTGGATG
CTGCTTCCCTGCCTCATTCTCCTGTGTCAGGTTCAAGGTGAAGAAACCCAGAAGGAACGTCC
CTCTCCACGGATCAGCTGTCCCAAAGGCTCCAAGGCCTATGGCTCCCCCTGCTATGCCTTGT
TTTGTCAACAAAATCCTGGATGGATGCAGATCTGGCTTGCCAGAACGGCCCTCTGGAAAA
CTGGTGTCTGTGCTCAGTGGGCTGAGGGATCCTCGTGTCCCTGGTGAAGGAGCATTAG
TAACAGCTACTCATACATCTGGATTGGGCTCCATGACCCCACACAGGGCTCTGAGCCTGATG
GAGATGGATGGGAGTGGAGTAGCACTGATGTGATGAATTACTTGCATGGAGAAAAATCCC
TCCACCATCTTAAACCCCTGGCCACTGTGGGAGCCTGTCAAGAACGACAGGATTCTGAAGTG
GAAAGATTATAACTGTGATGCAAAGTTACCCATGTCTGCAAGTTCAAGGACT**TAG**GGCAGGT
GGGAAGTCAGCAGCCTCAGCTGGCGTGCAGCTCATCATGGACATGAGACCAGTGTGAAGAC
TCACCCCTGGAAGAGAATATTCTCCCCAAACTGCCCTACCTGACTACCTTGTATGATCCTCC
TTCTTTCTTTCTTCACCTTCATTCAAGGCTTTCTGTCTTCCATGTCTTGAGATC
TCAGAGAATAATAATAAAAAATGTTACTTTATAAAAAAAAAAAAAAA

783/237

FIGURE 183

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56965
<subunit 1 of 1, 175 aa, 1 stop
<MW: 19330, pI: 7.25, NX(S/T): 1
MLPPMALPSVSWMLLSCLILLCQVQGEETQKELPSPRISCPKGSKAYGSPCYALFLSPKSWM
DADLACQKRPSGKLVSVLSGAEGSFVSSLVRSISNSYSYIWIGLHDPTQGSEPDGDGWEWSS
TDVMNYFAWEKNPSTILNPGHCGSLSRSTGFLWKWDYNCDAKLPYVCKFKD

Important features:**Signal peptide:**

amino acids 1-26

C-type lectin domain signature.

amino acids 146-171

784/237

FIGURE 184

CCAGTCTGTCGCCACCTCACTTGGTGTCTGCTGCCCCGCCAGGCAAGCCTGGGTGAGAGC
ACAGAGGAGTGGGCCGGACCATGCGGGGGACGCGGCTGGCGCTCCTGGCGCTGGTGTGGC
TGCCTGCGGAGAGCTGGGCCGGCCCTGCGCTGCTACGTCTGTCCGGAGCCCACAGGAGTGT
CGGACTGTGTCACCATGCCACCTGCACCAACGAAACCATGTGCAAGACCACACTCTAC
TCCCAGGAGATAGTGTACCCCTCCAGGGGGACTCCACGGTGACCAAGTCCTGTGCCAGCAA
GTGTAAGCCCTCGGATGTGGATGGCATGGCCAGACCCCTGCCGTGCTGCAATACTG
AGCTGTGCAATGTAGACGGGGCGCCGCTCTGAACAGCCTCCACTGCCGGCCCTCACGCTC
CTCCCACCTTGAGCCTCCGACTGTAGAGTCCCCGCCACCCCATGCCCTATGCCCTATGCCGCCA
GCCCGAATGCCCTGAAGAAGTGCCCTGCACCAAGGAAAAAAAAAAAAAAA

7851237

FIGURE 185

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56405
<subunit 1 of 1, 125 aa, 1 stop
<MW: 13115, pI: 5.90, NX(S/T): 1
MRGTRLALLALVLAACGELAPALRCYVCPEPTGVSDCVTIATCTTNETMCKTTLYSREIVYP
FQGDSTVTKSCASKCKPSDVGDIGQTLPVSCCNTELCNVDGAPALNSLHCGALTLLPLLSLRL

Important features:**Signal peptide:**

amino acids 1-17

N-glycosylation site.

amino acids 46-49

7861237

FIGURE 186

CTGCAGTCAGGACTCTGGGACCGCAGGGGGCTCCCGACCCCTGACTCTGCAGCCGAACCGGC
ACGGTTCTGGGGACCCAGGCTTGCAAAGTGACGGTCATTTCTCTTCTTCTCCCTCTT
GAGTCCTTCTGAGATGATGGCTCTGGCGCAGCGGGAGCTACCCGGGTCTTGTGCGATGG
TAGCGCGGCTCTGGCGGCCACCCTCTGCTGGGAGTGAGCGCCACCTGAACCTGGTTCTC
AATTCCAACGCTATCAAGAACCTGCCCGGACCGCTGGCGCGCTGCCGGCACCCAGGCTC
TGCAGTCAGCGCCGCCGGGAATCCTGTACCCGGCGGGATAAGTACCAAGACCATTGACA
ACTACCAGCCGTACCCGTGCGCAGAGGACGAGGAGTGCGGCACTGATGAGTACTGCGCTAGT
CCCACCCGGAGGGGACGCAGGCCTGCAAATCTGTCTGCCCTGCAGGAAGCGCCGAAAAGC
CTGCATGCCACGCTATGTGCTGCCCGGGATTACTGCAAAATGGAATATGTGTGCTT
CTGATCAAAATCATTCCGAGGAGAAATTGAGGAAACCATCACTGAAAGCTTGGTAATGAT
CATAGCACCTTGGATGGTATTCCAGAAGAACCACTTGTCTCAAAAATGTATCACACCAA
AGGACAAGAAGGTTCTGTTGCTCCGGTCATCAGACTGTGCCTCAGGATTGTTGTGCTA
GACACTTCTGGTCCAAGATCTGTAACCTGTCCTGAAAGAAGGTCAAGTGTGTACCAAGCAT
AGGAGAAAAGGCTCTCATGGACTAGAAATATTCCAGCGTTACTGTTGAGAAGGTCTGTC
TTGCCGGATACAGAAAGATCACCATCAAGCCAGTAATTCTCTAGGCTTACACTGTCAGA
GACACTAAACCAGCTATCAAATGCAGTGAACCTCTTATATAATAGATGCTATGAAAACC
TTTATGACCTTCATCAACTCAATCCTAAGGATATACAAGTTCTGTGGTTTAGTTAACGAT
TCCAATAACACCTCCAAAAACCTGGAGTGTAAGAGCTTGTCTTATGGAACCTCCCTG
TGATTGCAGTAAATTACTGTATTGTAATTCTCAGTGTGGCACTTACCTGAAATGCAATGA
AACTTTAATTATTTCTAAAGGTGCTGCACTGCCTATTTCTTGTATGTAATTT
TTGTACACATTGATTGTTACTGACTGACAAATATTCTATATTGAACTGAAGTAAATCATT
TCAGCTTATAGTTCTAAAGCATAACCTTACCCATTAAATTCTAGAGTCTAGAACGCA
AGGATCTCTGGAATGACAAATGATAGGTACCTAAATGTAACATGAAAATACTAGCTTATT
TTCTGAAATGTAATCTTAATGCTTAAATTATATTCCCTTCTAGGCTGTGATAGTTTGAA
AATAAAATTAAACATTAAAAAAAAAAAAAA

787/237

FIGURE 187

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57530
<subunit 1 of 1, 266 aa, 1 stop
<MW: 28672, pI: 8.85, NX(S/T): 1
MMALGAAGATRVFVAMVAAALGGHPLLGVSATLNSVLNSNAIKNLPPPLGGAAGHPGSAVSA
APGILYPGGNKYQTIDNYQPYPCAEDEECGTDEYCASPTRGGDAGVQICLACRKRRKRCMRH
AMCCPGNYCKNGICVSSDQNHFGRGEIEETITESFGNDHSTLDGYSRRTTLSSKMYHTKGQEG
SVCLRSSDCASGLCCARHFWSKICKPVLKEGQVCTKHRRKGSHGLEIFQRCYCCEGLSCRIQ
KDHHQASNSSLHTCQRH

Important features:**Signal peptide:**

amino acids 1-23

N-glycosylation site.

amino acids 256-259

Fungal Zn(2)-Cys(6) binuclear cluster domain

amino acids 110-126

781237

FIGURE 188

TGTGTTCCCTGCAGTCAGAATTGGGACNGCAGGGTTCCCGGACCTGATTTGCAGCGGA
ACGGGAAGGTTTGTGGGACCCAGGTTGAAATGACGGTCATTTTTCTTCCTTCNG
GAGTCCTNTGAGANGATGGTTTGGCGCAGCGGGAGCTAACCGGTTTGTNGCGATG
GTAGCGCGGTTTCCGGCGGCCACCTNTGCTGGGAGTGAGCGCCACCTTGAATCGGTTTC
AATTCCAACGNTATCAAGAACCTGCCACCGNTGGCGCGCTGGGGCACCCAGGNTT
TGCAGTCAGCGCCGCCGGAAATCCTGTACCCGGGGAAATAAGTACCAGACCATTGACA
ATTACCAAGCGTACCGTGCAGAGGACGAGGAGTGCGGACTGATGAGTACTGCGCTAGT
CCCACCCGCGGAGGGGANGCGGGCGTCAAATNTGTNTNGCCTGCAGGAAGCGCCAAAACG
CTGCATGCGTCANGCTATGTGCTGCCCGGGATTACTGCAAAATGGAATATGTGTNTT
CTGATCAAAATCATTCCGAGGAGAAATTGAGGAAACCATCACTGAAAGCTTGGTAATGAT
CATAGCACCTTGGATGG

789(237)

FIGURE 189

GAGGAACCTACCGGTACCGGCCGCGCGTGGTAGTCGCCGGTGGCTGCACCTCACCAATCCGTGCGCCGCG
 CTGGGCCGTCGGAGACTGCGTGTGCTCTCTCCTGCACGCGGTGTTGGCTCGGCCAGGCAGGGTCCGCCGCA
 GGGTTGAGGATGGGGAGTAGCTACAGGAAGCGACCCCGCATGGCAAGGTATTTTGTAATGCTGATTACCC
 AGTATTAGAAATGAGCTGAAGACCATTACAGATTAAATTTGGGACAGATTGTGATGCTGATTACCC
 TGAAGTAATGTAGACAGAAGTCTCAAATTTCATATTACACTGAACAGCAGTGAATCTTAATGTCAC
 TTAAATCAGAACATTGCAAGAAAGAATGGGAGTCTGGTAAATAAGTACTATCAGAGACTATGAAA
 GATCATTCTGTTCTGATAGTGTATGGCATTTTAGTGGGACAGATCAGGATTTCAGTGTACTGG
 AGTGTCCAAAATGCAAGCAGTAGAGAAATAAGACAAGCTTCAAGAAATTGGCATTGAGTTACATCCTGATA
 AAACCGAATAACCCAAATGCACATGGCATTTTAAAATAATAGAGCATATGAAGTACTCAAAGATGAAG
 TCTACGGAAAAAGTATGACAATATGGAGAAAAGGGACTTGAGGATAATCAAGGTGCCAGTATGAAAGCTGAA
 CTATTATCGTTATGATTTGGTATTATGATGATGATGATCCTGAAATCATAACATTGAAAAGAAGAGAATTGATGC
 TGCTGTTAATCTGGAGAACTGTGGTTGTAAATTACTCCCAGGCTGTTACACTGCCATGATTTAGCTCC
 CACATGGAGAGACTTTGCTAAAGAAGTGGATGGTTACTCGAATTGGAGCTGTTACTGTGGTGTGATGAGAAT
 GCTTGGCAATGAAAGGAGTCAACAGCTATCCAGTCCTCATTTGGTGTGAAATGGCCCCAGTGAATA
 TCATGGAGACAGTCAAAGGAGAGTTAGTGTGATTGCAATGCAAGCATGTTAGAAGTACAGTGACAGAACTTG
 GACAGGAAATTGTCACACTCAAACGTCTTGTGCTGGTATTGGCTGGTGTACTTTGTCAAA
 AGGAGGAGATTGTTGACTTCACAGACACGACTCAGGCTTAGTGGCATGTTGTTCTCAACTCATTGGATGCTAA
 AGAAATATATTGGAAGTAATAACATAATCTCCAGATTGAACTACTTCGGCAAACACACTAGAGGATGTT
 GGCTCATCATCGGTGGTGTATTTTCAATTGGAAAAATGAAATGATCCTGAGCTGAAAAAAACT
 AAAACTCTACTTAAATGATCATTCAGTGGCAGGTTGACTGTTCTGACAGACATCTGAGTAA
 TCTGATGTTTCAGCCGCTCTAGCAGTATTAAAGGACAAGGAAACAAAGAATGAAATTCTCATGAA
 GAAGATTCTATGATATACTTGCTTGCCTTGCAGAAAGTGTGAAATTCTCATGTTACACGCTGGACCTCAAA
 TTTCTGCCATGACAAAGAACATGGCTGTTGATTCTTGCCCCCTGGTGTGCCACATGTCAGCTTACT
 ACCAGAGTTACGAAGAGCATCAAATCTCTTATGGTCAGCTTAAGTTGGTACACTAGATTGACAGTTCATGA
 GGGACTCTGTAACATGATAACATTCAAGCTTATCCAACACAGTGGTATTCAACCAGTCCAACATTGAGTA
 TGAGGACATCACTCTGTAACAAATCTGGAGGTTCATAGAGGATCTTATGAAATCCTCAGTGGCTCCCTAC
 ACCACCCACCTCAACGAACTAGTACACAAAGAAAACACAACGAAAGTGGATGGTGTGATTTCTATTCTCGTG
 GTGTCATCTGCAAGTCTTAATGCCAGAATGAAAGAACATGGCCCCGACATTAACTGGACTGATCAACGTGG
 CAGTATAGATTGCCAACAGTATCATTCTTGTGCCCAGGAAACGTTCAAGATAACCTGAGATAAGATT
 TCCCCAAAATCAAATAAGTTACAGTATCACAGTTACATGGTGGATAGGGATGCTTATCCCTGAGAAT
 CTGGGGCTAGGATTTACCTCAAGTATCCACAGATCTAACACCTCAGACTTCACTGAAAAAGTTCTACAAGG
 GAAAATCTGGGTGATTGATTCTATGCTCCTGGTGTGACCTTGCAGAATTGCTCCAGAATTGAGCT
 CTTGGCTAGGATGATTAAAGGAAAGTGAAGGAAAGCTGGAAAAGTAGACTGTCAGGCTTATGTCAGACATGCCAGAA
 AGCTGGGATCAGGGCTATCCAACGTTAACAGTAAAGGAGCAGAACAGGAAATTCTAAGAAGAGCA
 GATAAAATACAGAGATGCAAAGAACATCGCTGCCCTAATAGTGAAGGAAACTCTCCGAAATCAAGGCAA
 GAGGAAATAAGGATGAACTTGTAAATGAGATGAAACATTATCTAGACTTGCAAGTGTACTGCCA
 GAATTATCTACAGCACTGGTAAAAGAAGGGCTGCAAACATTCTGTAAGGGCCGTTATAAAATT
 GACTTGCAGGCTATAATATGGTCACACATGAGAACAGAACAGTCATCATGTTGTTATTGCT
 TTTAACACCTTAAAAAATTTAACGATTCTAGCTCAGAGCCATACAAAGTAGGCTGGATTGAGTCATG
 GACCATGATTGCTGTCCTCGACGGACTTAAATGTTCAAGGCTGGCTGACATGAGTCTGCTG
 ATCTACATAATGTCTAAGTTGATAAAAGTCCACCTTCCCTCACGTTGGCTGACCTGAAAAGAGGTA
 ACTTAGTTGGTCACTGTTCTCTAAATGCTATCCCTAACATATTATATTCTGTTAAAGGAGGTT
 GATGTGGCACAGTAAACAAACCTGTTATGCTGATTATTATGAGGAGATTCTCATTGTTCTCTCTCA
 AAGGTTGAAAAAATGCTTAAATTTTCAACAGCCGAGAACAGTGCAGCAGTATATGTCACACAGTAAGTACAC
 AAATTGAGCAACAGTAAGTGCACAAATTCTGTAGTTGCTGATCATCCAGGAAACCTGAGGGAAAAAATT
 TAGCAATTAACTGGCATTGAGTATCTAAATATGTTATCAAGTATTAGAGTTCTATATTAAAGATATA
 TGTGTTCATGATTCTGAAATTGCTTCAAGGAAATTCTCCTGACTGATAGTTGATTTGAGGCATCTA
 TTACATATTGCTTCTGAACTTGTGACCTGATCTTCTTATTACATTGGGTTTCTTCAAGTTG
 TTTTCACTCTGTCAGTCTATTATTAACTGAAAGGAAATTACTTACAGGTTGTTTACTGTAGCTT
 AATGATACTGTAGTTATTCCAGTTACTAGTTACTGTCAAGGGCTGCCCTTTCAGATAAATATTGACATAATA
 ACTGAAGTTATTGTTATAAGAAAATCAAGTATATAATCTAGGAAAGGGATCTCTAGTTCTGTTGTTAGA
 CTCAAAGAATCACAAATTGTCAGTAACATGTAGTTGTTAGTTATAATTCAAGGTACAGAATGGTAAAATT
 CCAATCAGTCAAAGAGGTCATGAATTAAAGGCTGCAACTTTCAAAAAAAAAAAAAA

7901237

FIGURE 190

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56439
<subunit 1 of 1, 747 aa, 1 stop
<MW: 86127, pI: 7.46, NX(S/T): 2
MGVWLNKDDYIRDLKRIILCFLIVYMAILVGTQDFYSLLGVSKTASSREIRQAFKKLALKL
HPDKNPNNPAHGDFLKINRAYEVLKDEDLRKKYDKYGEKGLDNQGGQYESWNYYRYDFGI
YDDDPEIITLERREFDAAVNSGELWFVNFYSPGCSHCHDLAPTRDFAKEVDGLLRIGAVNC
GDDRMLCRMKGVNNSYPSLFIFRSGMAPVKYHGDRSKESLVSFAMQHVRSTVTELWTGNFVNS
IQTAFAAGIGWLITFCSKGGDC LTSQTRLRLSGMLFLNSDAKEIYLEVIHNLPDFELLSAN
TLEDRLAHHRWLLFFHFGKNENSNDPELKKLKTLKNDHIQVGRFDSSAPDICSNLYVFQP
SLAVFKGQGTKEYEIHHGKKILYDILAFAKESVNHSVTTLGPQNFPANDKEPWLVDFFAPWC
PPCRALLPELRRASNLLYGQLKFGTLDCTVHEGLCNMYNIQAYPTTVVFNQSNIHEYEGHHS
AEQILEFIEDLMNPSVSVSLTPPTFNEVLTQRKHNEVMVDFYSPWCHPCQVLMPEWKRMART
LTGLINVGSIDCQQYHSFCAQENVQRYPEIRFFPPKSNKAYQYHSYNGWNRDAYSLRIWGLG
FLPQVSTDLPQTFSKVLQGKNHWVIDFYAPWCGPCQNFAPFELLARMIKGKVAKGVDC
QAYAQTCKAGIRAYPTVKFYFYERAKRNFOEEQINTRDAKAIAALISEKLETLRNQGKRNKDEL

Important features:**Endoplasmic reticulum targeting sequence.**

amino acids 744-747

Cytochrome c family heme-binding site signature.

amino acids 158-163

Nt-dnaJ domain signature.

amino acids 77-96

N-glycosylation site.

amino acids 484-487

7911237

FIGURE 191

AGACAGTACCTCCTCCCTAGGACTACACAAGGACTGAACCAGAAGGAAGGAGACAGAGCAAA
GCCATGAACATCATCCTAGAAATCCTCTGTTCTGATCACCATCATCTACTCCTACTTGGA
GTCGTTGGTGAAGTTTCATTCTCAGAGGAGAAAATCTGTGGCTGGGAGATTGTTCTCA
TTACTGGAGCTGGGCATGGAATAGGCAGGCAGACTACTTATGAATTGCAAAACGACAGAGC
ATATTGGTTCTGTGGATATTAATAAGCGCGGTGGAGGAAACTGCAGCTGAGTGCGAAGA
ACTAGGCCTCACTGCGCATGCGTATGTGGTAGACTGCAGCAACAGAGAAGAGATCTATCGCT
CTCTAAATCAGGTGAAGAAAGAAGTGGGTGATGTAACAATCGTGGTGAATAATGCTGGACA
GTATATCCAGCCGATCTCAGCACCAAGGATGAAGAGATTACCAAGACATTGAGGTCAA
CATCCTAGGACATTTGGATCACAAAAGCACTTCTCCATCGATGATGGAGAGAAATCATG
GCCACATCGTCACAGTGGCTCAGTGTGCGGCCACGAAGGGATTCTACCTCATCCCATA
TGTTCAGCAAATTGCCGCTGTTGGCTTCACAGAGGTCTGACATCAGAACTTCAGGCCT
GGGAAAAACTGGTATCAAAACCTCATGTCCTGCCCAGTTTGTGAATACTGGGTTACCA
AAAATCCAAGCACAAGATTATGGCCTGTATTGGAGACAGATGAAGTCGTAAGAAGTCTGATA
GATGGAATACTTACCAATAAGAAAATGATTTGTTCCATCGTATATCAATATCTTCTGAG
ACTACAGAAGTTCTTCTGAACGCGCCTCAGCGATTTAAATCGTATGCAGAAATATTCAAT
TTGAAGCAGTGGTTGCCACAAAATGAAATTGAATAAGCTCCAGGCAGAGATG
TATGCATGATAATGATATGAATAGTTCGAATCAATGCTGCAAAGCTTATTCACATT
TCAGTCCTGATAATATTAACATTGGTTGGCACTAGCAGCAGTCAAACGAACAAGATTA
ATTACCTGTCTCCTGTTCTCAAGAATATTCAGTAGTTTCTAGGTCTGTTTCCT
TCATGCCTCTAAAAACTCTGTGCTTACATAAACATACTTAAAGGTTTCTTAAGATAT
TTTATTTCCATTAAAGGTGGACAAAAGCTACCTCCCTAAAGTAAATACAAAGAGAACT
TATTTACACAGGGAAAGGTTAACAGACTGTCAAGTAGCATTCAATCTGTAGCCATGCCACAG
AATATCAACAAGAACACAGAACAGTGCACAGCTAACAGAGATCAAGTTCAGCAGGCAGCTT
ATCTCAACCTGGACATATTTAACAGATTCAAGCATTGAAAGATTCCCTAGCCTCTCCTT
TCATTAGCCCCAAACGGTGCAACTCTATTCTGGACTTATTACTTGATTCTGTCTGTAT
AACTCTGAAGTCCACCAAAAGTGGACCCCTCTATATTCCTCCCTTTATAGTCTTATAAGA
TACATTATGAAAGGTGACCGACTCTATTTAAATCTCAGAACATTAAAGTTCTAGCCCCATGA
TAACCTTTCTTGTAAATTATGCTTCATATATCCTGGTCCCAGAGATGTTAGACAAT
TTTAGGCTAAAAATTAAAGCTAACACAGGAAAAGGAACGTACTGGCTATTACATAAGAAA
CAATGGACCCAAAGAGAAAGAA

792/237

FIGURE 192

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56409
<subunit 1 of 1, 300 aa, 1 stop
<MW: 33655, pI: 9.31, NX(S/T): 1
MNIILEILLLITIIYSYLESVKFFIPQRRKSVAGEIVLITGAGHGIGRQTTYEFAKRQSI
LVLWDINKRGVEETAAECRKLGVTAHAYVVDCSNREIYRSLNQVKKEVGDVTIVVNNAGTV
YPADLLSTKDEEITKTFEVNILGHFWITKALLPSMMERNHGHIVTVASVCVGHEGIPYLIPYC
SSKFAAVGFHRGLTSELQALGKTGIKSCLCPVFVNTGFTKNPSTRLWPVLETDEVVRSЛИD
GILTNKKMIFVPSYINIFRLQKFLPERASAILNRMQNIQFEAVVGHKIKMK
```

Important features:**Signal peptide:**

amino acids 1-19

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 30-33 and 58-61**Short-chain alcohol dehydrogenase family protein**
amino acids 165-202, 37-49, 112-122 and 210-219

793/237

FIGURE 193

CGGC GG CGG CT GCGGG CGC GAGGT GAGGGG CGC GAGGT GAGGGG CGC GAGGT TCCC AGC AGG
ATGCC CGG CT TG CAGGA AGCT GAAGT GAGAGGCC GGAGAGGGCC CAGCC CGCC GGGC
AGGATG ACCA AGGCC CGG CT GTTCC CGG CT GTGG CT GGT GCT GGGG TCGG TGT CATG AT CCT
GCT GATCAT CGT TACT GGG ACAG CGC AGG CGCC CGC ACT TCA TT GAC ACAGT CCT TCT
CTAGG CGC ACAC GGGG CC CGC CT GCC AC GCC CGG ACAGGG ACAGGG AGCT CAC G
GCC GACT CCG AT GT CGAC GAGTT CT GGAC AAGTT CT CAGT GCT GG CGT GAAG CAG AGCGA
CCTTCC CAGAAAGG AGAC GGAG CAGCC GCT GCG CC GGG GAGC AT GGAGG AGAG CGT GAGAG
GCTAC GACT GGT CCC CGC GAC GCCC GGCG CAGCC CAG ACCA GGGG CGG CAG CAGG CGG AG
CGGAGG AGCGT GCT GCGGG CT TCT GCGCC AACT CCAGC CT GGC CT TCCC ACCA AGG AGCG
CGCATT CGAC GAC AT CCCC AACT CGG AGCT GAGCC ACCT GAT CGT GGAC GAC CGG CAC GGG
CCAT CT ACT GCT AC GTGCC AAGGT GG CTC GACCA ACT GGA AGCG CGT GAT GAT CGT GCT G
AGCGGA AGC CT GCT GCA CC CGG TGC CCGT ACC CG GAC CC GCT GCG C AT CCC CGC GAGCA
CGT GCA CAACGCC AGCG CGCAC CT GAC CT CAACA AGTT CT GGCG CGCT AC GGG AAGC TCT
CCC GCC ACCT CAT GAAGGT CAAGCT CAAGA AGT AC ACCA AGTT CCT TCGT GCG CGAC CCCC
TTC GTG CGC CT GAT CT CGC CT TCG CAGCA AGTT CGAG CT GGAGA AC GAGG AGTT CT ACCG
CAAGTT CGCC GTGCC AT GCT GCGG CT GTAC GCA ACCA ACCA GCG CT GCC CGCT CGC G
GCGAGG CTT CGC GTGG CTC AAGGT GTCC TCG CCA ACT CAGT ACC GCT GCG GAC
CCGC ACACGG AGAG CAGT GGC GCC CTT CAAC GAG CAGT GGC GGG CAGG TGT ACC GCG CT GCG
CCC GTGCC AGAT CGACT AC GACT TCGT GGGGA AGCT GGAG ACT CT GGAC GAGG AC GCG CGC
AGCT GCT GCA GCT ACT CCAGG TGG ACC GGC AGC TCC GCT TCCCCC GAGC TAC CGG AAGC AGG
ACCGCC AGC AGCT GGGAGG AGG ACT GTG TCG CCA AGAT CCCC CT GGCG CT GGAGG CAGC AGC
GTATAAA ACT CTAC GAGG CGA CTT GTT CT CTT CGC TAC CCA AGC CCG AAA ACCT CCT CC
GAGA CTGAA AGCT TCG CGT TGT GCT TTT CT CGC GTG CCGT GGA ACCT GAC GCA CGC GCA CTCC
AGTTTTT ATGAC CTAC GATT TGT GCA AT CTGG CTT CT TGT TCA CTCC ACT GCG CT TATCC
ATTGAGT ACT GTAT CGA TATT GTTTTA AGATTA ATAT ATT CAGG TATT AATACGA

794/237

FIGURE 194

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56112
<subunit 1 of 1, 414 aa, 1 stop
<MW: 48414, pI: 9.54, NX(S/T): 4
MTKARLFRLWLVLGSVFMILLIIVYWDSAGAAHFYLHTSFSRPHTGPPPLPTPGPD RDRELT A
DSDVDEFLDKFLSAGVKQSDLPRKETEQPPAPGSMEESVRGYDWSPRDARRSPDQGRQQAER
RSVLRGFCANSSLAFPTKERAFFDI PNSELSHLIVDDRHGAIYCYVPKVACTNWKRVMIVLS
GSLLHRGAPYRDPLRIPREHVHNASAHLTFNKFWR RY GKL SRH LMKV KLY TKFL FVRDPF
VRLISAFRSKFELENEEFYRKFAVPMLRLYANHTSLPASAREAFRAGLKVSFANFIQYLLDP
HTEKLAPFNEHWRQVYRLCHPCQIDYDFVGKLETLD EDAQLLQ LLQ VDRQLRFPPSYRNRT
ASSWEEDWFAKIPLAWRQQLYKLYEADFVLFGYPKPENLLRD

Important features:**Signal peptide:**

amino acids 1-31

N-glycosylation sites.

amino acids 134-137, 209-212, 280-283 and 370-373

TNFR/NGFR family cysteine-rich region protein

amino acids 329-332

795/237

FIGURE 195

TCGGGCCAGAATTCCGGCACGAGGCACGAGGGCGACGGCCTCACGGGCTTGAGGTGA
AAGAGGCCAGAGTAGAGAGAGAGAGACCGACGTACACGGG**ATGG**GCTACGGAACCGCCT
ATGCCGGAAAGGTGGTGGTCGTGACCGGGGCGGGCGGCATCGGAGCTGGGATCGTGC
GCCTTCGTGAACAGCGGGGCCGAGTGGTTATCTGCACAAGGATGAGTCTGGGGCCGGC
CCTGGAGCAGGAGCTCCCTGGAGCTGTCTTATCCTCTGTGATGTGACTCAGGAAGATGATG
TGAAGACCTGGTTCTGAGACCATCCGCCATTGGCCCTGGATTGTGTTGTCAACAAC
GCTGGCCACCACCCACCCCCACAGAGGCCTGAGGAGACCTCTGCCAGGGATTCCGCCAGCT
GCTGGAGCTGAACCTACTGGGACGTACACCTGACCAAGCTGCCCTCCCTACCTGCCGA
AGAGTCAAGGGAAATGTCATCAACATCTCCAGCCTGGTGGGGCAATCGGCCAGGCCAGGCA
GTTCCCATATGGTGTCCGAGTCAACTGTATCTCCCAGGAAACATCTGGACCCGCTGTGGG
AGGAGCTGGCAGCCTTAATGCCAGACCCCTAGGCCACAATCCGAGAGGGCATGCTGCCAG
CCACTGGGCCGCATGGGCCAGCCCGCTGAGGTGGGGCTGCCAGTGTCCCTGCCCTCCGA
AGCCAACCTCTGCACGGCATTGAACTGCTCGTGACGGGGGTGCAGAGCTGGGTACGGGT
GCAAGGCCAGTCGGAGCACCCCCGTGGACGCCCGATATCCCT**TGA**TTCTCTCATTT
CTACTTGGGGCCCCCTCCTAGGACTCTCCCACCCAAACTCCAACCTGTATCAGATGCAGC
CCCCAAGCCCTTAGACTCTAACGCCAGTTAGCAAGGTGCCGGTCACCCCTGCAGGTTCCCAT
AAAAACGATTGCAGCC

796/237

FIGURE 196

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56045
<subunit 1 of 1, 270 aa, 1 stop
<MW: 28317, pI: 6.00, NX(S/T): 1
MATGTRYAGKVVVVTGGGRGIGAGIVRAFVNNSARVVICDKDESGGRALEQELPGAVFILCD
VTQEDDVKTLVSETIRRGRLDCVVNNAGHHPPPQRPEETSQGFRQLLELNLLGTYTLKL
ALPYLRKSQGNVINISSLVGAIGQAQAVPVYATKGAVTAMTKALALDESPYGVRCNCISPGN
IWTPLWEELAALMPDPRATIREGMLAQPLGRMGQPAEVGAAAVFLASEANFCTGIELLVTGG
AELGYGCKASRSTPVDAPDIPS
```

Important features:**N-glycosylation site.**

amino acids 138-141

Short-chain alcohol dehydrogenase family protein

amino acids 10-22, 81-91, 134-171 and 176-185

797/237

FIGURE 197

AGGCGGGCAGCAGCTGCAGGCTGACCTTGCAGCTGGCGGA**ATG**ACTGGCCTCACAAACCTG
CTGTTCTTCTTACCATTTCCATCTCCTGGGCTGGGCCAGCCCAGGAGCCCCAAAGCAA
GAGGAAGGGCAAGGGCGGCCTGGGCCCTGGCCCTGGCCCTCACCAAGGTGCCACTGGACC
TGGTGTACGGATGAAACCGTATGCCGCATGGAGGAGTATGAGAGGAACATCGAGGAGATG
GTGGCCAGCTGAGGAACAGCTCAGAGCTGGCCAGAGAAAGTGTGAGGTCAACTTGCAGCT
GTGGATGTCCAACAAGAGGAGCCTGTCTCCCTGGGCTACAGCATCAACCACGACCCAGCC
GTATCCCCGTGGACCTGCCGGAGGCACGGTGCCTGTGTCTGGGCTGTGAACCCCTTCACC
ATGCAGGAGGACCGCAGCATGGTGAGCGTGCCGGTGTTCAGCCAGGTTCTGTGCGCCGCCG
CCTCTGCCGCCACCGCCCCGCACAGGGCCTGCCGCCAGCGCGCAGTCATGGAGACCATCG
CTGTGGGCTGCACCTGCATCTTCT**TGA**ATCACCTGGCCAGAAGCCAGGCCAGCAGCCCCGAGA
CCATCCTCCTTGCACCTTGTGCCAAGAAAGGCCTATGAAAAGTAAACACTGACTTTGAAA
GCAAG

7981237

FIGURE 198

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA59294
<subunit 1 of 1, 180 aa, 1 stop
<MW: 20437, pI: 9.58, NX(S/T): 1
MDWPHNLLFLLTISIFLGLGQPRSPKSKRKGQGRPGPLAPGPHQVPLDLVSRMKPYARMEY
ERNIEEMVAQLRNSSELAQRKCEVNLQLWMSNKRSLSPWGYSINHDPSRIPVVDLPEARCLCL
GCVNPFTMQEDRSMVSVPVFSQVPVRRRLCPPPRTGPCRQRAVMETIAVGCTCIF
```

Important features:**Signal peptide:**

amino acids 1-20

N-glycosylation site.

amino acids 75-78

Homologous region to IL-17

amino acids 96-180.

7991237

FIGURE 199

GC GCCGCCAGGCGTAGGC GG GTGGCC CTT GCGT CCCC GCTT CTT GAAAA ACCGGCGGG
CGAGCGAGGCTGC GG CG CTGCC CTT CCCC AC ACT CCCC GCGAGAACGCTCGCTCG
GCGCCAACATGGCGGGTGGCGCTGC GG CCGCAGCTAACGGCGCTCCTGGCGCCTGGAT
CGCGGCTGTGGCGGCACGGCAGGCCGGAGGAGGCGCTGCCGCCGGAGCAGAGCCGG
TCCAGCCC ATGACC GCGCTCCA ACTGGACGCTGGTGTGGAGGGCGAGTGGATGCTGAAATT
TACGCCCCATGGTGTCCATCCTGCCAGCAGACTGATTCAAGAATGGAGGCTTTGCAAAGAA
TGGTGAATACTTCAGATCAGTGTGGGAAGGTAGATGTCATTCAAGAACAGGTTGAGTG
GCCGCTCTTGTCACC ACTCTCCAGCATT TTT CATGCAAAGGATGGATATTCCGCGT
TATCGTGGCCCAGGAATCTCGAAGACCTGCAGAATTATATCTTAGAGAACAGAACATGGCAATC
AGTCGAGCCTCTGACTGGCTGGAAATCCCCAGCTCTCTAACGATGTCTGGAAATGGCTGGTC
TTTTAGCATCTCGCAAGATATGGCATCTTCACA ACTATTTCACAGTGACTCTGGAAATT
CCTGCTGGTGTCTTATGTGT TTT CGTCATAGCCACCTGGTTTGCCTTTATGGG
TCTGGTCTGGTGTGGTAATATCAGAATGTTCTATGTGCCACTTCAAGGCAATTATCTGAGC
GTTCTGAGCAGAACATGGAGATCAGAGGAGGCTCATAGAGCTGAACAGTTGCAAGGATGCAG
GAGGAAAAAGATGATTCAAATGAAGAACAAAAGACAGCCTGTAGATGATGAAGAAGA
GAAAGAACATCTGGCGATGAGGATGAAGCAGAGGAAGAGGAGGAGGACAACCTGGCTG
CTGGTGTGGATGAGGAGAGAACGTGAGGCCATGATCAGGGGCCAGGAGGACGGTGTG
ACCCGGGAGGAAGTAGAGCCTGAGGAGGCTGAAGAACGGCATCTGAGCAACCTGCCAGC
TGACACAGAGGTGGTGAAGACTCCTTGAGGCAGCGTAAAGTCAGCATGCTGACAAGGGAC
TGTAGATTTAATGATGCGTTCAAGAACACACCAAAACATATGTCAGCTCCCTTGG
CCTGCAGTTGTACCAAATCCTTAATT TCTGAATGAGCAAGCTCTCTTAAAGATGCT
CTCTAGTCATTGGTCTCATGGCAGTAAGCCTCATGTATACTAAGGAGAGTCTCCAGGTGT
GACAATCAGGATATAGAAAAACAAACGTAGTGTGGATCTGTTGGAGACTGGATGGAA
CAAGTTCATTACTTAGGGTCAGAGAGTCTGACCAGAGGAGGCCATTCCAGTCCTAATC
AGCACCTCCAGAGACAAGGCTGCAGGCCCTGTGAAATGAAAGCCAAGCAGGAGCCTGGCT
CCTGAGCATCCCCAAAGTGTAACTAGAAGCCTTGCACTCTTCTGTGAAAGTATTAT
TTTGTCAAATTGCAGGAAACATCAGGCACCACAGTGCATGAAAATCTTCACAGCTAGAA
ATTGAAAGGCCTGGGTATAGAGAGCAGCTCAGAAGTCATCCCAGCCCTGTGAATCTCCTG
TGCTATGTTATTCTTACCTTAATT TCCAGCATTCCACCATGGGCATTAGGCTCT
CCACACTCTCACTATTATCTTGGTCAGAGGACTCCAATAACAGCCAGGTTACATGAAC
TGTGTTGTTCAATTCTGACCTAAGGGTTAGATAATCAGTAACCATAACCCCTGAAGCTGT
GACTGCCAACATCTCAAATGAAATGTTGTGGCATCAGAGACTCAAAGGAAGTAAGGATT
TTACAAGACAGATTAAGGAAATTGTTGTCCAAATATAGTTGTGTTGATTTTTTT
AAGTTTCTAAGCAATATTTCAGGCCAGAACGCTCTAACGTTGCTGAGTACAAGGTAGT
CTTGTGAAGAAAAGTTGAATACTGTTGTGTTCATCTCAAGGGTCCCTGGTCTTGAAC
TACCTTAATAATAACTAAAAACCACTCTGATTTCCTCAGTGTGCTTTGGTGA
GAATTAAATGAACCTCCAGTACCTGAAAGTGAAAGATTGATTTGTTCCATCTGTAATC
TTCCAAAGAATTATCTTGTAAATCTCTCAATACTCAATCTACTGTAAGTACCCAGGGAG
GCTAATTCTT

200/237

FIGURE 200

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56433
<subunit 1 of 1, 349 aa, 1 stop
<MW: 38952, pI: 4.34, NX(S/T): 1
MAGGRCGPQLTALLAAWIAAVAAATAGPEEAALPPEQSRVQPMTASNWTLVMEGEWMKFYAP
WCPSCQQTDSEWEAFAKNGEILQISVGKVDVIQEPEGLSGRFFVTLPAFFHAKDGIFRRYRG
PGIFEDLQNYILEKKWQSVEPLTGWKSPASLTMSGMAGLFSISGKIWHLHNYFTVTLGIPAW
CSYVFFVIATLVFGLFMGLVLVVISECFYVPLPRHLSERSEQNRRSEEAHRAEQLQDAEEEK
DDSNEEENKDSLVDDEEEKEDLGDEDEAEEEEEDNLAAGVDEERSEANDQGPPGEDGVTR
EVEPEEAEGISEQPCPADTEVVVEDSLRQRKSQHADKGL
```

Important features:

Signal peptide:

amino acids 1-22

Transmembrane domain:

amino acids 191-211

N-glycosylation site.

amino acids 46-49

Thioredoxin family proteins. (homologous region to disulfide isomerase)

amino acids 56-72

Flavodoxin proteins

amino acids 173-187

207/237

FIGURE 201

ATCTGGTTGAAC TAACTTAAGCTTAACTTGTAAACTCCGGTAAGTACCTAGCCCACATGATT
TGACTCAGAGATTCTCTTTGTCCACAGACAGTCATCTCAGGGGCAGAAAAGAGCTCC
CAAATGCTATATCTATTCAAGGGCTCTCAAGAACAATGGAATATCATCCTGATTAGAAAAT
TTGGATGAAGATGGATATACTCAATTACACTTCGACTCTCAAAGCAATACCAGGATAGCTGT
TGTTCAGAGAAAGGATCGTGTGCTGCATCTCCTCCTGGCGCTCATTGCTGTAAATTG
GAATCCTATGCTGGTAATACTGGTGTAGCTGTGGTCTGGGTACCATGGGGTTCTTCC
AGCCCTGTCTCCTAATTGGATTATATGAGAAGAGCTGTTATCTATTCAGCATGTCACT
AAATTCTGGGATGGAAGTAAAAGACAATGCTGGCAACTGGCTCTAATCTCCTAAAGATAG
ACAGCTCAAATGAATTGGGATTATAGTAAAACAAGTGTCTCCCAACCTGATAATTCAATT
TGGATAGGCCTTCTCGGCCCGACTGAGGTACCATGGCTCTGGGAGGATGGATCAACATT
CTCTCTAACTTATTCAGATCAGAACACAGCTACCCAAGAAAACCCATCTCAAATTGTG
TATGGATTCACTGTCAGTCATTATGACCAACTGTGTAGTGTGCCCTCATATAGTATTG
GAGAAGAAGTTTCAATGTAAGAGGAAGGGTGGAGAAGGGAGAGAAAATATGTGAGGTAGTA
AGGAGGACAGAAAACAGAACAGAAAAGAGTAACAGCTGAGGTCAAGATAATGCAGAAAATG
TTTAGAGAGCTGGCCAAC TGTAATCTAACCAAGAAATTGAAGGGAGAGGCTGTGATTCT
GTATTGTCGACCTACAGGTAGGCTAGTATTATTTCTAGTTAGTAGATCCCTAGACATGG
AATCAGGGCAGCCAAGCTTGAGTTTATTATTATTATTGAGATAGGGTCT
CACTTGTACCCAGGCTGGAGTGCAGTGGCACATCTGACTCACTGCAGCTATCTCTCGC
CTCAGCCCCCAAGTAGCTGGACTACAGGTGCATGCCACCATGCCAGGCTAATTGGT
TTTTTGTAGAGACTGGTTTGCCATGTTGACCAAGCTGGTCTCTAACCTGGCTTAAG
TGATCTGCCGCCTGGCCTCCAAAGTGTGGATTACAGATGTGAGCCACACCTGGC
CCCAAGCTTGAATTTCATTCTGCCATTGACTGGCATTAACCTGGTAAAGCCATAAGCGA
ATCTTAATTCTGGCTCTATCAGAGTTGTTCATGCTCAACAATGCCATTGAAGTGCACGGT
GTGTTGCCACGATTGACCCCTCAACTCTAGCAGTATATCAGTTATGAACTGAGGTGAAAT
ATATTCTGAATAGCTAAATGAAGAAATGGGAAAAAAATCTCACCACAGTCAGAGCAATT
ATTATTTCATCAGTATGATCATAATTGATTATCATCTTAGTAAAGCAGGAACCTCTA
CTTTTCTTATCAATTAAATAGCTCAGAGACTACATCTGCCATATCTCTAACATAGAATCTT
TTTTTTTTTTTTTGAGACAGAGTTCGCTCTTGTGCCCCAGGCTGGAGTGCAACGG
CACGATCTCGGCTACCGAACCTCCGCCCTGGGTTCAAGCAATTCTCCTGCCCTAGCCT
CCCAAGTAGCTGGGATTACAGTCAGGCACCACACCCGGCTAATTGTATTAGT
AGAGACAGGGTTCTCCATGTCGGTCAGGGTAGTCCGACTCTGACCTCAAGTGTACTGC
CTGCCTCGGCCTCCAAAGTGTGGATTACAGGCGTGAGCCACTGCACCCAGCCTAGAATCT
TGTATAATATGTAATTGTAGGAAACTGCTCTCATAGGAAAGTTCTGCTTTAAATACA
AAAATACATAAAATACATAAAATCTGATGATGAATATAAAAAGTAACCAACCTCATTGGA
ACAAGTATTAAACATTGGAAATATGTTATTAGTTGTGATGTACTGTTACAATT
ACCATTTCAGTAATTACTGAAAATGGTATTATTGGAAATGAAACTATAATTCCCTCATG
TGCTGATTGTCTTATTGTCATACTTCCCACCTGGTGCTATTGTTATTCCAATGGATA
TTCTGTATTACTAGGGAGGCATTACAGTCCTCTAATGTTGATTAATATGTGAAAAGAAAT
TGTACCAATTACTAAATTATGCACTTAAAGGGATGATTATGTTATGTTATGTTATG
TTCAATAAAAAAAAAACTCTTATCAAAAAAAAAAAAAAA

202 / 237

FIGURE 202

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53912
<subunit 1 of 1, 201 aa, 1 stop
<MW: 22563, pI: 4.87, NX(S/T): 1
MEYHPDLENLDEDGYTQLHFDSQSNTRIAVVSEKGSCAASPPWRLIAVILGILCLVILVIAV
VLGTMGVLSGPCPPNIIYEKSCYLFSMSLNSWDGSKRQCWQLGSNLLKIDSSNELGFIVKQ
VSSQPDNSFWIGLSRPQTEVPWLWEDGSTFSSNLFQIRTTATQENPSPNCVWIHVSVIYDQL
CSVPSYSICEKKFSM

Important features:**Type II transmembrane domain:**

amino acids 45-65

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 197-200**N-myristoylation sites.**

amino acids 35-40 and 151-156

Homologous region to LDL receptor

amino acids 34-67 and 70-200.

2031237

FIGURE 203

204/237

FIGURE 204

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50921
<subunit 1 of 1, 693 aa, 1 stop
<MW: 77738, pI: 8.87, NX(S/T): 7
MTPQSLLQTLFLLSLLFLVQGAHGRGHREDFRFCQRNQTHRSSLHYKPTPDLRISIENSE
EALTVHAPFPAAHPASRSFPDPRLGLYHFCLYWNRHAGRLHLLYGRDFLLSDKASSLLCFQH
QEEESLAQGPPLLATSVTSWWSPQNISLPSAASFTFSFHSPPHTAAHNASVDMCELKRDLQLL
SQFLKHPQKASRRPSAAPASQQLQSLLESKLTsvRFMGDMVSFEEDRINATVWKLQPTAGLQD
LHIHSRQEEEQSEIMEYSVLLPRTLFQRTKGRSGEAEKRLLLVDFSSQALFQDKNNSQVLGE
KVLGIVVQNTKVANLTEPVVLTQFHQLQPKNVTLQCVFWEDPTLSSPGHWSSAGCETVRRE
TQTSCFCNHLYFAVLMVSSVEVAHKHYLSLLSYVGCVVSALACLVTIAAYLCSRVPPLC
RRKPRDYTIKVHMNLLAVFLLDTSFLLSEPVALTGSEAGCRASAIFLHFSLLTCLSWMGLE
GYNLYRLVVEVFGTYVPGYLLKLSAMGWGFPIFLVTLVALVDVDNYGPIILAVHRTPEGVIY
PSMCWIRDSDLVSYITNLGLFSLVFLFNMAMLATMVVQILRLRPHTQKWSHVLTLLGLSLVLG
LPWALIFFSFASGTQQLVVLYLFSIITSFQGFLIFIWYWSMRLQARGGPSPLKSNSDSARLP
ISSGSTSSSR

Important features:**Signal peptide:**

amino acids 1-25

Putative transmembrane domains:amino acids 382-398, 402-420, 445-468, 473-491, 519-537, 568-590
and 634-657**Microbodies C-terminal targeting signal.**

amino acids 691-693

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 198-201 and 370-373

N-glycosylation sites.amino acids 39-42, 148-151, 171-174, 234-237, 303-306, 324-327
and 341-344**G-protein coupled receptors family 2 proteins**

amino acids 475-504

2051237

FIGURE 205

TGCCTGGCCTGCCCTGTCAACAATGCCGTTACTCTGCTTCCAGGTTGCCCTGCCCTGCAGA
GGAAANCNTCGGGACTACACCNTCAAGTCACATGAACCTGCTGCTGCCGTCTCCTGCTG
GACACGAGCTTCCCTGCTCAGCGNAGCCGGTGGCCCTGACAGGCTCTGAAGGCTGGCTGCCGA
GCCAGTGCCATCTTCCCTGCACTTCTCCTGCTCACCTGCCTTCCCTGGATGGGCCTCGAGGGG
TACAACCTCTACCGACTCGTGGTGGAGGTCTTGGCACCTATGTCCCTGGCTACCTACTCAA
GCTGAGCGCCATGGGCTGGGCTTCCCCATCTTCTGGTACGCTGGTGGCCCTGGTGGATG
TGGACAACTATGGCCCCATCATCTTGGCTGTGCATAGGACTCCAGAGGGCGTCATCTACCC
TCCATGTGCTGGATCCGGACTCCCTGGTCAGCTACATACCAACCTGGGCCTTTCAGCCT
GGTGTTCCTGTTCAACATGG

206/237

FIGURE 206

CGGACGCGTGGCGGACGCGTGGCGGACGCCTGGCGTGGCTGGTCAGGTCCAGGTTGCTTG
TCCTTTCAAAAATGGAGACACAGAAGAGGGCTCTAGGAAAAAGTTGGATGGATTATGTGGAAACTACCC
GCGATTCTCTGCTGCCAGACAGCAGGCTGGCGCTTCCACCCAGTCAGCCTCCCTGGCGGTGGTGAAGAGAC
TCGGGAGTCGCTGCTTCCAAAGTGGCCCGCGTGAATGAGCTCTCACCCAGTCAGCCAAATGAGCCTCTCGGGC
TTCTCCTGCTGACATCTGCCCTGGCCCGAGAGACGGGACTCAGGGCAATCAGGAACTTACAGCTGAGTAGTAAATTCC
AGTTTCCAGCAACAAGGAACAGAACGGAGTACAAGATCCTCAGCATGAGAGAAATTACTGTGCTACTAATG
GAAGTATTACAGCCAAAGGTTCTCATACTTATCCAAGAAATACGGCTTGGTATGGAGATTAGTAGCAGTAG
AGGAAAATGTATGGATAACAACCTACGTTGATGAAAGAGATTGGGCTTGAAGACCCAGAAGAGATGACATATGCAAGT
ATGATTTGAGAAGTTGAGGAACCCAGTGATGGAACATATTAGGGCCTGGTGTGGTCTGGTACTGTACCAAG
GAAAACAGATTCTAAAGGAATCAAATTAGGATAAGAGATTGATCTGATGAATATTTCCTCTGAACCAGGGT
TCTGCATCCACTACAACATGTGATGCCAACTCACAGAAGCTGTGAGTCCTCAGTGTACCCCTTCAGCCT
TGCCACTGGACCTGCTTAATATGCTATAACTGCCTTAGTACCTGGAGACCTTTCGATATCTTGAACCCAG
AGAGATGGCAGTTGGACTTAAAGATCTATATAGGCAACTTGGCAACTCTTGGCAAGGCTTTGGGAA
GAAAATCCAGAGTGGGATCTGAACCTCTAACAGAGGAGGTAAGATTACAGCTGCACACCTCGTAACCTCT
CAGTGTCCATAAGGAAGAACTAAAGAGAACCGATACCATTCTGCCAGGTTGTCCTCTGGTAAACGCTGT
GTGGGAACGTGCTGTTGTCTCCACAATTGCAATGAATGTCAATGTGCTCCAGCAGAAAGTTACTAAAAAATACC
ACGAGGTCTTCAGTTGAGACCAAAGACCGGTGTCAGGGGATTGCACAAATCACTCACCGACGTGGCCCTGGAGC
ACCATGAGGAGTGTGACTGTGTCAGAGGAGCACAGGAGGATAGCCGCATCACCAACCGAGCAGCTCTGCCCA
GAGCTGTGAGTGGCAGTGGCTGATTCTATTAGAGAACGTTGCTTATCCTTAATCTCAGTTGCT
TCAAGGACCTTCATCTCAGGATTACAGTCAGTGCATTCTGAAAGAGGAGACATCAAACAGAATTAGGAGTTGCT
ACAGCTCTTGAGAGGAGGCTAAAGGACAGGGAGAAAAGGTTCTCAATGTTGAAAGAAAATTAAATGTTGCT
TAAATAGATCACCAGCTAGTTTCAAGATTACATGTCAGTACAGGAAAAAAACTGTGCAAGTGAGCACCTGATTCGTTGCT
GATACTGGCTTAGGTAATGTCAGTACAGGAAAAAAACTGTGCAAGTGAGCACCTGATTCGTTGCT
TCTAAAGCTCCATGTCCTGGCCTAAATCGTATAAAATCTGGATTTTTTTTTTGCTCATATTACAT
ATGTAACACAGAACATTCTAATGACTACAAACCTGGTTTAAAGGAACATGTTGCTATGAATTAAACTGT
GTCATGCTGATAGGACAGACTGGATTTCATATTCTTATTAAATTCGCAATTAGAAGAAGAGAACTACA
TTCATGGTTGGAGAGATAAAACCTGAAAAGAGTGGCCTATCTTCACTTTATCGATAAGTCAGTTATTTG
TTTCATTGTCATTTTATATTCTCTTGTGACATTAAACTGTGCTTTCTAATCTGTTAAATATATCT
ATTTTACCAAAGGTATTAATATTCTTTATGACAACCTAGATCAACTATTTAGCTGGTAAATTTCT
AAACACAATTGTTATAGCCAGAGGAACAAAGATGATATAAAATTGTTGCTCTGACAAAAATACATGTATTCA
TTCTCGTATGGTCTAGAGTTAGATTAATCTGCATTAAACTGTAATTGGAATAGAATTGGTAAGTTGCAAA
GACTTTGAAAATAATTAAATTATCATATCTTCATTCCTGTTATTGGAGATGAAAATAAAAAGCAACTTATGA
AAGTAGACATTCAAGTCCAGCCATTACTAACCTATTCTTTGGGGAAATCTGAGCCTAGCTCAGAAAAACAT
AAAGCACCTGAAAAGACTTGGCAGCTCTGATAAGCGTGTGCTGAGTAGGAACACATCCTATTAA
TTGTGATGTTGGTTTATTATCTAAACTCTGTCATACACTGTATAAAATACATGGATATTGTTATGTAC
GAAGTATGTCCTAACCAAGTTCACTTATTGACTCTGGCAATTAAAAGAAAATCGTAAATATTGCT
AAAATGCTTAATATGTCCTAGGTTATGTGGTACTATTGAATCAAAATGTATTGAATCATCAAATAAAAGA
ATGTGGCTATTGGGGAGAAAATTAAAAAAAAAAAGGTTAGGGATAACAGGGTAATGCGGCC

207/237

FIGURE 207

MSLFGLLLTSALAGQRQGTQAESNLSSKFQFSSNKEQNGVQDPQHERIITVSTNGSIHSPR
FPHTYPRNTVLVWRLVAVEENVWIQLTFDERFGLEDPEDDICKYDFVEVEEPSDGTILGRWC
GSGTVPKGKQISKGNQIRIRFVSDEYFPSEPGFCIHYNIVMPQFTEAVSPSVLPPSALPLDLL
NNAITAFSTLEDLIRYLEPERWQLDLEDLYRPTWQLLGKAFVGRKSRVVDLNLLTEEVRLY
SCTPRNFSVSIREELKRTDTIFWPGCLLVKRCGGNCACCLHNCNECQCVP SKVTKKYHEVLO
LRPKTGVRGLHKSLTDVALEHHEECDCVCRGSTGG

Signal sequence:

amino acids 1-14

20/237

FIGURE 208

CCCATCTCAAGCTGATCTGGCACCTCTCATGCTCTGCTCTTCAACCAGACCTACATTCCATTGGAAAGA
 AGACTAAAAATGGTTTTCCAATGTGGACACTGAAGAGACAAATTCTATCCTTTAACATAATCCTAATTCC
 AAACCTCTGGGCTAGATGGTTCTAAAACCTCTGCCCTGTGATGTCACTCTGGATGTTCAAAGAACCATGTG
 ATCGTGGACTGCACAGACAAGCATTGACAGAAATTCTGGAGGTATTCCCACGAACACCACGAACCTCACCTC
 ACCATTAACCACATACCAGACATCTCCCCAGCCTTCAGACTGGACCATCTGGTAGAGATCGATTTCAGA
 TGCAACTGTGACCTATTCCACTGGGGTCAAAAAACAAACATGTCATCAAGAGGCTGAGATTAAACCCAGAAGC
 TTTAGTGGACTCACTTATTAAATCCTTACCTGGATGAAACCCAGCTACTAGAGATACCGCAGGGCCTCCG
 CCTAGCTACAGCTCTCAGCCTGAGGCCAACAAACATCTTCCATCAGAAAAGAGAATCTAACAGAACTGGCC
 AACATAGAAATACTCTACCTGGGCCAAACTGTTATTATCGAAATCCTGTTATGTTATTCATATTCAATAGAGAAA
 GATGCCCTCTAAACTTGACAAAGTTAAAGTGTCTCCCTGAAAGATAACAATGTCACAGCCGTCCACTGTT
 TTGCCATCTACTTTAACAGAACTATATCTACAAACAACATGATTGCAAAAATCCAAGAAGATGATTTAATAAC
 CTCAACCAATTACAAATTCTTGACCTAAAGTGGAAATTGCCCCATTCTGTCGCGCC
 TGTTAAAATATTCTCCCTACAGATCCCTGTAATGCTTTGATGCGCTGACAGAATTTAACGATCTGCTAA
 CACAGTAACTCTTCAGCATGTGCCCAAGATGGTTAAAGAACATCAACAAACTCCAGGAACGGATCTGTC
 CAAAACCTCTGGCCAAGAAATTGGGATGCTAAATTCTGCATTTCTCCCAGCCTCATCCAATTGGATCTG
 TCTTCATTTGAACCTCAGGTCTATCGTCATCTATGAATCTATCACAAGCATTCTTCACTGAAAAGCCTG
 AAAATTCTGCGATCAGAGGATATGCTTTAAAGAGTTGAAAAGCTTTAACCTCTGCCATTACATAATCTCAA
 AATCTTGAAAGTCTTGATCTGGCACTAACTTTATAAAAATGCTAACCTCAGCATGTTAAACAAATTAAAAGA
 CTGAAAGTCATAGATCTTCAGTGAATAAAATCACCTCAGGAGATTCAAGTGAAGTTGGCTCTGCTCAAAT
 GCCAGAACCTCTGAGAAAGTTGACACCCAGGCTTGAACAAATTACATATTTCAGATATGATAAGTATGCA
 AGGAGTTGCGAGATTCAAAAACAAAGAGGCTTCTTCATGTCGTTAATGAAAGCTGCTACAAGTATGGCAGACC
 TTGGATCTAAGTAAAATAGTATATTGTCAGTCTCTGATTTCAGCATCTTCTTCAAATGCC
 AATCTGTCAGGAAATCTCATTAGCCAAACTCTTAATGGCAGTGAATTCCAACCTTTAGCAGAGCTGAGATATTG
 GACTCTCCAACAACCGGCTTGAATTACTCCATTCAACAGCATTGAAAGAGCTTCACAAACTGGAAGTTCTGGAT
 ATAAGCAGTAATAGCCATTATTTCAATCAGAAGGAATTACTCATATGCTAAACTTACCAAGAACCTAAAGGTT
 CTGAGAACCTGAGATGACATGACATCTCTCCACAGCAGGAGCATGGAGACTGAGTCTCTTAGA
 ACTCTGGAATTCTGAGGAAACTCTTGTGATGTTATGGAGAGAAGGTGATAACAGATACTTACAAATTATTCAG
 AATCTGCTAAAATTAGAGGAATTAGACATCTCTAAATTCCCTAAGTTCTGCTTCTGGAGTTTGATGGT
 ATGCCTCAAATCTAAAGAATCTCTTTGGCCAAAATGGCTCAAATCTTCAGTTGGAAGAAACTCCAGTGT
 CTAAGAACCTGGAAACTTGGACCTCAGCCACAACCAACTGACCAACTGCCCTGAGAGATTATCAAACACTGTTCC
 AGAACGCTCAAGAATCTGATTCTTAAGAATAATCAAACATCAGGAGTCTGACGAAGTATTTCTACAGATGCCCTC
 CAGTGGCGATATCTGGATCTCAGCTCAAATAAAATCCAGATGATCCAAAAGACCAGCTTCCCAGAAAATGTCCTC
 AACAAATCTGAAGATGTTGCTTGCATCATAATCGGTTCTGTCACCTGTGAGTGTGTTGGCTGCTGGGG
 GTTAACCATACGGAGGTGACTATTCTTACCTGGCACAGATGTCAGTTGTTGGCCAGGAGCACACAAGGGC
 CAAAGTGTGATCTCCCTGGATCTGACACCTGTGAGTTAGATCTGACTAACCTGATTCTGTTCTACTTCCATA
 TCTGTATCTCTTTCTCATGGTGTGATGACAGCAAGTCACCTCTATTCTGGGATGTGTTGATATTACCAT
 TTCTGTAAGGCCAAGATAAAGGGGTATCAGCGCTAATATCACCAGACTGTTGCTATGATGCTTTATTGTGAT
 GACACTAAAGACCCAGCTGTGACCGAGTGGGTTTGGCTGAGCTGGTGGCCAAACTGGAAGACCAAGAGAGAAA
 CATTAAATTATGTCGAGGAAAGGGACTGGTTACCAAGGGCAGCCAGTTCTGGAAAACCTTCCCAGAGCATA
 CAGCTTAGCAAAAAGACAGTGTGATGACAGACAAGTGTGAAAGACTGAAATTTAAGATAGCATTTCAC
 TTGTCCTCATCAGAGGCTATGGATGAAAAGTTGATGTCGTTGATGTTCTGAGTGGCCACAAACCCGCAAGCTCAC
 CCATACTCTGGCAGTGTCTAAAGAACGCCCTGGCCACAGACAAATCATGTCAGGCTATAGTCAGGTGTTCAAGGAA
 ACGGTCTAGCCCTTCTTGCAAAACACAACACTGCCTAGTTACCAAGGAGAGGCTGGC

209/237

FIGURE 209

MVFPMWTLKRQILILFNIILISKLLGARWFPKTLPCDVTLDVPKNHVIVDCTDKHLTEIPGG
IPTNTTNLTILTINHIPDISPASFHRLDHVEIDFRNCVPIPLGSKNNMCIKRLQIKPRSFS
GLTYLKSLYLDGNQLEIPQGLPPSLQLLSLEANNIFSIRKENTELANIEILYLGQNCYR
NPCYVSYSIEKDAFLNLTKLKVSLKDNNVTAVPTVLPSLTLYLYNNMIAKIQEDDFNNL
NQLQILDLSGNCPRCYNAPFPCAPCKNNSPLOIPVNNAFDALTELKVLRLHSNSLQHVPPRF
KNINKLQELDLSQNFLAKEIGDAKFLHFLPSLIQLDLSFNFELQVYRASMNLSQLAFSSLKSL
KILRIRGYVFKEKSFNLSPLHNLQNLEVLDLGTNFKIANLSMFQFKRLKVIDLSVNKIS
PSGDSSEVGFCNSARTSVESYEPOVLEQLHYFRYDKYARSCRFKNKEASFMSVNESCYKGQ
TLDLSKNSIFFVKSSDFQHLSFLKCLNLSGNLISQTLNGSEFQPLAELRYLDFSNRLLH
STAFAEELHKLEVLDISSLNSHYFQSEGITHMLNFTKNLKVLQKLMMNDNDISSTSRTMESES
LRTLEFRGNHLDVLWREGDNRYLQLFKNLLKLEELDISKNSLSFLPSGVFDGMPPNLKNL
AKNGLKSFSWKKLQCLKNLETLDLSHNQLTTPERLSNCRSRSLKNLILKNNQIRSLTKYFLQ
DAFQLRYLDLSSNKIQMIQKTSFPENVLNNLKMLLLHHNRFLCTDAWFVWWVNHTEVТИР
YLATDVTCVGPAGHKGQSVISLDLYTCEDLTNLILFSLISISVSLFLIMVMMTASHLYFWDVW
YIHFCKAKIKGYQRLISPDCYDAFIIVYDTKDPAVTEWVLAELVAKLEDPREKHFNLCEE
RDWLPGPVLENLSQSICLSSKTVFVMTDKYAKTENFKIAFYLSHQRLMDEKVDVIILIFL
KPFOQSKFLQLRKRLCGSSVLEWPTNPQAHPYFWQCLKNALATDNHVAYSQVFKETV

Signal sequence:

amino acids 1-26

Transmembrane domain:

amino acids 840-860

210 / 237

FIGURE 210

GGGTACCATTCTGCGCTGTCGAAGTACGGAATGAAAAATTAGAACACAGAAACATGGAAAACATGTTCCCTC
 AGTCGTCAATGTCGACCTGCATTTCTGCTAATATCTGTTCTGTGAGTTATGCCCGAAGAAAATTTC
 GAAGCTATCCTGTGATGAGAAAAGCAAATGACTCAGTTATTGCAGAGTGCAGCAATCGTCACTACAGGAAG
 TTCCCCAAACGGTGGCAAAATATGTGACAGAACTAGACCTGCTGATAATTCACTCACACACATAACGAATGAAT
 CATTCAAGGGCTGCAAAATCTCACTAAAATAACTAAACACACCAATGTACAGCACCAGAACCGAAATC
 CGGTATAACAAATGCTGAAATCATCACAGACGGGCACTCTCAACCTAAACCTAAGGGAGTTACTGC
 TTGAAGAACACCGATTACCCCAAATACCCCTCTGGTTGCCAGAGTCTTGACAGAACTTAGTCTAATTCAAAC
 ATATATACAACATAACTAAAGAGGGCATTCAAGACTTAAACTTGAAAATCTCTATTGGCTGGAACGT
 ATTAAACAAAGTTGCGAGAAAACTAACATAGAAGATGGAGTATTGAAACGCTGACAATTTGGAGTTGCTAT
 CACTATCTTCACACAGTGCACCCAAACTGCCAAGCTCCCTACGCAAACATTCTGAGCAACA
 CCCAGATCAAATACATTAGTAAGAAGATTCAAGGGATTGATAAATTAAACATTACTAGATTAAAGCGGGAACT
 GTCGAGGTGCTTCAATGCCCAATTCCATGCGTGCCTGTGATGGTGGTCTCAATTAAATATAGATGTTTG
 CTTCAAAATCTGACCCAACTTCGATACTAAACCTCTCTAGCACTTCCCTCAGGAGAGAAATGCTGCCTGGT
 TAAAGGATATGCTCATCTGAAGGTGCTGGATCTGAATTCAACTATTAGTGGGAGAAATAGTCTGGGGCAT
 TTTAACGATGTCGCCCCGTTAGAAAATCTGACTGTCTTTAACTATATAAGGGAGTTATCCACAGCATA
 TTAATATTCCAGAAACTTCTCTAAACTTTGTCCTACGGGCAATTGCAATTAAAGAGGTTATGTTCCAGGAAC
 TCAGAGAAGATGATTCCAGCCCTGATGCAGCTCCAAACTTATGACTATCAACTTGGTATAATTATTAA
 AGCAAATCGATTCAAACCTTCCAAATTCTCCAATCTGAAATTATTACTTACTGTCAGAAAACAGAATATCAC
 CGTGGTAAAGATAACCGGCAGAGTATGCAAATAGTCTCTTTCAAGTCATATCGGAAACGACGCTCAA
 CAGATTGAGTTGACCCACATTGCACTTTATCATTCCACCCGCTTTAAATAAGCCACAAATGTCGTT
 ATGAAAAGCCTAGATTAAAGCCTAACAGTATTCTCATTGCCCCAAATTGAAAATCTCTGACA
 TTGCTGTTAAATCTGCTGCAATAGCAATGCTCAAGTGTAAAGGAACTGAAATTTCAGCATTCCCTCATG
 TCAAATATTGGATTGACAAACAATAGACTAGACTTGTATAATGCTAGTGCTTACTGAATTGTCGACTGG
 AAGTCTAGATCTCAGCTATAATTCAACTATTTCAGAATAGCAGCGTAACACATCATCTAGAATTATTCAA
 ATTTCACAAATCAAAGTTAAACTGAGGCCAACACATTATACTTTAACAGATAAGTATAACCTGGAA
 GCAAGTCCCTGGTAGAATTAGTTTCAGTGGCAATGCCCTGACATTGTTGGAATGATGATGACAACAGGTATA
 TCTCCATTTCAGGCTCAAGAACATGACACGCTGCTGGATTATCCCTTAATAGGCTGAAGCACATCCCAAATG
 AAGCATTCTTAATTGCCAGCGAGTCTCACTGAACATATAAAATGATAATATGTTAAAGGTTAACTGGA
 CATTACTCCAGCAGTTCTCGTCTGAGTTGACTTACGGAAACAAACTACTCTTTAACTGATAGCC
 TATCTGACTTTACATCTCCCTCGGACACTGCTGACTGTCATAACAGGATTCCACCTACCCCTGGCTTC
 TTTCTGAAGTCAGTAGTCTGAAGCACCTCGATTAAAGTCCAATCTGCTAAAACAATCAACAAATCCGCACTTG
 AAACGTTAACAGGACCACCAATTATCTATGTTGAAACTACACGGAAACCCCTTGAATGCACTGTGACATTGGAG
 ATTTCGAGATGGATGGATGAAACATCTGAATGTCAAAATTCCAGACTGGTAGATGTCATTGTCAGTCACTG
 GGGATCAAAGGGAGAGTATTGAGCTGGAGCTAACAACTTGTGTTAGTCTGAGTCACTGCACTGATATT
 TTTCTCACGTTCTTATCACCACCATGGTTATGGCTGCCCTGGCTCACCATTGTTACTGGGATGTT
 GGTTATATATAATGTTGTTAGCTAAAGGCTACAGGCTCTTCCACATCCCAAACCTTCTATGATG
 CTTACATTCTTATGACACCAAAGATGCCCTGTTACTGACTGGGTGATAATGAGCTGCGTACCCATTGAG
 AGAGCGAGACAAAAGCTCTCCTTGTCTAGAGGAGAGGGATTGGGACCCGGATTGCCATATCGACAACC
 TCATGCAGAGCATCAACCAAAGCAAGAAAACAGTATTGTTAACAAAAATATGCAAAAGCTGAACTTAA
 AACAGCTTTTACTTGGCTTGCAGAGGCTAATGGATGAGAACATGGATGTGATTATATTCTGCTGGAGC
 CAGTGTACAGCATTCTCAGATTGGAGCTACGGCAGGGATCTGTAAGAGCTCCATCCAGTGGCAGTCA
 ACCCGAAGGCGAGGCTTGGCAAAACTCTGAGAAATGTTGACTGAAATGATTCAACGGTATAACA
 ATATGTTATGTCGATTCAAGCAATACTACTGACGTTAACGTCTGAGTCACTGTTGGCTCATAATAAGATGCAAG
 GAATGACATTCTGTTAGTTATCTATTGCTATGTAACAAATTATCCAAAACATTAGTGGTTAAAACAACACA
 TTTGCTGGCCCACAGTTTGAGGGTCAGGAGTCCAGGCCAGCATAACTGGGCTCTGCTCAGGGTGTCTCAG
 AGGCTGCAATGTTAGGCTTACCAAGAGACATAGGCATCACTGGGTCACACTCATGTTGTTGTTCTGGATTCA
 ATTCCCTGGCTATTGGCAAAGGCTATACTCATGTAAGCCATGCGAGCCCTCCACAAGGCAGCTGCTTC
 ATCAGAGCTAGCAAAAGAGAGGGTTGCTAGCAAGATGAACTCACAAATCTTGTAAATGCAATCAAAGGAGT
 ATCTCATCATTGGCCATATTCTATTGTTAGAAGTAAACCAACAGGCTCCACAGCTCATGGGATGACCA
 TCAGTCCAGGGAAAACAGCTGAAAGACAGATGGTGGAGCTGATTGCTTCAGTTGGTCACTCAACTATTCC
 TGACTGCTGCTGGGATGCCCTGCTATTGTTGAGTATGTTGAAATATCAGGAGGGAGGATCACTGTGGACC
 ATCTTAGCAGTTGACCTAACACATCTTCTTTCAATATCTAAGAACATTGCACTGTGACTAATGGTCTAATA
 TTAAGCTGTTTATTTATCATATATCTATGCTACATGTTATATTGCTGTTGCTGTTGCTGGTTTAT
 TTACAGTTGCTTTACAAATATTGCTGAAACATTGACTTCTAAGGTTAGATGCCATTAAAGAACTGAGAGTGG
 ATAGCTTTAAAGCATTTTACTTACCAATTAAAGTATGCAAGCTAAATTGCAAGCTTTGGTCTATA
 TTGTTAATTGCCATTGCTGAAATCTAAAATGAATGAATAAAATGTTCAATTACAAAAAA

211/237

FIGURE 211

MENMFLQSSMLTCIFLLISGSCELCAEENFSRSYPCDEKKQNDSVIAECSNRRLQEVPQTVG
KYVTELDLSDNFITHITNESFQGLQNLTKINLNHNPNVQHQNGNPGIQSNGLNITDGAFLNL
KNLRELLLEDNQLPQIPSGLPESLTELSSIQNNIYNITKEGISRLINLKNLYLAWNCYFNKV
CEKTNIEDGVFETLTNLELLSLSFNSLSHVPPKLPSSLRKLFLSNTQIKYISEEDFKGLINL
TLLDLSGNCPRCFNAPFPCVPCDGGASINIDRFAFQNLTQLRYLNLSSTSLRKINAAWFKNM
PHLKVLDEFNYLGVGEIVSGAFLTMLPRLIEILDLSFNYIKGSYPQHINISRNFSKLLSLRAL
HLRGYVFQELREDDFQPLMQLPNLSTINLGINFQKQIDFKLFQNFSNLEIIYLSNRISPLV
KDTRQSYANSSSFQRHIRKRRSTDFFDPHSNFYHFTRPLIKPQCAAYGKALDLSLNSIFFI
GPNQFENLPDIACLNLSANSNAQVLSGTEFSAIIPHVKYLDLTTNNRLDFDNASALTELDLEV
LDLSYNSHYFRIAGVTHHLEFIQNFTNLKVLNLSHNNIYTLDKYNLESKSLVELVFSGNRL
DILWNDDDNRYISIFKGLKNLTRLDSLNRKHIPNEAFLNLPASLTELHINDNMLKFFNWT
LLQQFPRLELLDLRGNKLLFLTDSDLSDFTSSLRTLLLSHNRISHLPSGFLSEVSSLKHLDLS
SNLLKTINKSALETKTTKLSMLEHGPNFECTCDIGDFRRWMDEHLNVKIPRLVDVICASP
GDQRGKSIVSLELTTCVSDVTAVILFFFITTMVMLAALAHHLFYWDVWFYIYNVCLAKVK
GYRSLSTSQTFYDAYISYDTKDASVTDWVINELRYHLEESRDKNVLLCLEERDWDPGLAIID
NLMQSINQSKKTVFVLTKKYAKSWNFKTAFYLAQRLMDENMDVIIFILLEPVLQHSQYLRL
RQRICKSSILQWPDPNPKAEGLFWQTLRNVVLTENDSRYNNMYVDSIKQY

Signal sequence:

amino acids 1-26

Transmembrane domain:

amino acids 826-848

212/237

FIGURE 212

CCAGGTCCAAC TGCACCTCGTTCTATCGATTGAATTCCCCGGGATCCTCTAGAGATCCCT
CGACCTCGACCCACCGTCCGCCAAGCTGGCCCTGCACGGCTGCAAGGGAGGCTCCTGTGGA
CAGGCCAGGCAGGTGGCCTCAGGAGGTGCCTCAGGCCAGTGGCCTGAGGCCAGC
AAGGGCTAGGGTCCATCTCCAGTCCCAGGACACAGCAGCGGCCACCATGGCCACGCCCTGGC
TCCAGCAGCATCAGCAGCCCCCAGGACCAGGGAGGCACAGGTGGCCCCCACCACCCGGAGGA
GCAGCTCCTGCCCTGTCCGGGGATGACTGATTCTCCTCGCCAGGCCACCCAGAGGAGAA
GGCCACCCCGCCTGGAGGCACAGGCC**ATG**AGGGCTCTCAGGAGGTGCTGCTGATGTGGCTT
CTGGTGTGGCAGTGGCGGCACAGAGCACGCCTACCGGCCGGCGTAGGGTGTGCTGT
CCGGGCTCACGGGACCCCTGTCTCGAGTCGTTCGTGACCGTGTTGACAGCCCTTCCTCA
CCACCTGCACGGGACCCGGCCTGCAGCACCTACCGAACCATCTATAGGACCGCCTACCGC
CGCAGCCCTGGGCTGGCCCTGCCAGGCCTCGCTACCGTGCTGCCCGGCTGGAAGAGGAC
CAGCGGGCTTCCCTGGGCCTGTGGAGCAGCAATATGCCAGCCGATGCCGGAACGGAGGGA
GCTGTGTCCAGCCTGGCGCTGCCGCTGCCCTGCAGGATGGCGGGTGACACTTGCCAGTCA
GATGTGGATGAATGCAGTGCTAGGAGGGCGGCTGTCCCCAGCGCTGCATCAACACCGCCGG
CAGTTACTGGTGCCAGTGTGGAGGGCACAGCCTGTCTGCAGACGGTACACTGTGTGC
CCAAGGGAGGGCCCCCAGGGTGGCCCCAACCCGACAGGAGTGGACAGTGCAATGAAGGAA
GAAGTGCAGAGGCTGCAGTCCAGGGTGGACCTGCTGGAGGAGAACGCTGCAGCTGGTGTGGC
CCCAC TGCAAGCCTGGCCTGCCAGGCACTGGAGCATGGCTCCCGAACCCGGCAGCCTCC
TGGTGCACTCCTTCCAGCAGCTGGCCCATCGACTCCCTGAGCGAGCAGATTCCCTCCTG
GAGGAGCAGCTGGGTCTGCTCTGCAAGAAAGACTCG**TGA**CTGCCAGCGCCCCAGGCTG
GACTGAGCCCTCACGCCGCCCTGCAGCCCCATGCCCTGCCAACATGCTGGGGTCCAG
AAGCCACCTGGGGTGA CTGAGCGGAAGGCCAGGCAGGGCCTCCTCCTCTTCCCTCCCC
TTCCTGGGAGGCTCCCCAGACCCCTGGCATGGATGGCTGGATCTCTGTGAATCCAC
CCCTGGCTACCCCCACCCCTGGCTACCCCAACGGCATCCAAAGGCCAGGTGGCCCTCAGCTG
AGGGAAAGGTACGAGCTCCCTGCTGGAGCCTGGACCCATGGCACAGGCCAGGCAGCCGGAG
GCTGGGTGGGGCTCAGTGGGGCTGCTGCCTGACCCCCAGCACAATAAAATGAAACGTGA
AAAAAAAAAAAAAAAAAAAAAAAAGGGCGGCCCGACTCTAGAGT
CGACCTGCAGAAGCTTGGCCGCCATGCCCAACTGTTATTGCAGCTATAATGGTTACAAAT

273/237

FIGURE 213

MRGSQEVL LMWLLV LAVGGTEHAYRPGR RVCAVRAHGD PVSE SFVQRVYQPFLTTCDGHRAC
STYRTIYRTAYRRSPGLAPARPRYACCPGWKRTSGLPGACGAAICQPPCRNGGSCVQPGRCR
CPAGWRGDTCQSDVDECSARRGGCPQRCINTAGSYWCQCWE GHSL SADGTL CVPKG GPPRVA
PNPTGVDSAMKEEVQRLQSRVDLLEEKIQLVLAPLHSLASQALEHGLPD PGSLLVHSFQQQLG
RIDSLSEQISFLEEQLGSCSCKKDS

Signal sequence:

1-19

214/237

FIGURE 214

GCCAGGCAGGTGGGCCTCAGGAGGTGCCCTCAGGCGGCCAGTGGGCCTGAGGGCCCCAGCAAG
GGCTAGGGTCCATCTCCAGTCCCAGGACACAGCAGCGGCCACCAGGCCACGCCCTGGGCTCC
AGCAGCATCAGAGCAGCCCCCTGTGGTTGGCAGCAAAGTTCAGCTTGCTGGCCCTGTGA
GGGGCTTCGCGCTACGCCCTGCGGTGTCCTGGAGGGCTGAGGTCTCCATCTTCTCCCTAGC
AGTGGATGAGCAACCCAACGGGGCCGGGGAGGGAACTGGCCCCAGGGAGAGGAACCCC
AAAGCCACATCTGTAGCCAGGATGAGCAGTGTGAATCCAGGCAGCCCCCAGGACCAGGGAGG
CACAGGTGGCCCCCACCACCCGGAGGAGCAGCTCCTGCCCTGTCCGGGGATGACTGATT
TCCTCCGCCAGGCCACCCAGAGGAGAAGGCCACCCGCCCTGGAGGCACAGGCCATGAGGGC
TCTCAGGAGGTGCTGCTGATGTGGCTCTGGTGTGGCAGTGGCAGGGCAGAGCACGCC
CCGGCCCGGGCGTAGGGTGTGCTGTCCGGCTCACGGGACCCCTGTCTCGAGTCGTT
TGCAGCGTGTGTACCAAGCCCTTCCCTACCCACCTGCGACGGGCACCGGGCCTGCAGCACCTAC
CGAACCATCTATAGGACCGCCTACCGCCGAGCCCTGGCTGGCCAGGAGCAGCAATAT
CGCGTGCTGCCCGGGCTGGAAGAGGACCAGCAGGGCTTCCCTGGGCCTGTGGAGCAGCAATAT
GCCAGCCGCATGCCGAACGGAGGGAGCTGTGCTGCCAGCCCTGGCGCTGCCCTGCA
GGATGGCGGGGTGACACTTGCCAGTCAGATGTGGATGAATGCAGTGCTAGGAGGGCGGCTG
TCCCCAGCGCTGCATCAACACCGCCGGCAGTTACTGGTGCCAGTGTGGAGGGCACAGCC
TGTCTGCAGACGGTACACTCTGTGTGCCAAGGGAGGGCCCCCAGGGTGGCCCCAACCCG
ACAGGAGTGGACAGTGCAATGAAGGAAGAAGTGCAGAGGCTGCAGTCCAGGGTGGACCTG
GGAGGAGAAGCTGCAGCTGGTGTGGCCACTGCACAGCCTGGCCTCGCAGGCAGTGGAGC
ATGGGCTCCCGGACCCGGCAGCCTCCTGGTGCACTCCTCCAGCAGCTGGCCGCATCGAC
TCCCTGAGCGAGCAGATTCCCTGGAGGAGCAGCTGGGTCTGCTCCTGCAAGAAAGA
CTTGATGCCTGCCAGCGCTCCAGGCTGGACTGAGCCCTCACGCCGCCCTGCAGCCCCCATG
CCCCTGCCAACATGCTGGGGTCCAGAAGCCACCTCGGGGTGACTGAGCGGAAGGCCAGGC
AGGGCCTTCCCTCTCCCTCCCTGGAGGCTCCCCAGACCCCTGGCATGGGAT
GGGCTGGATCTCTGTGAATCCACCCCTGGCTACCCCAACCCCTGGTACCCCAACGGCA
TCCCAAGGCCAGGTGGACCCCTCAGCTGAGGGAAAGGTACGAGCTCCCTGCTGGAGCCTGGAC
CCATGGCACAGGCCAGGCAGCCGGAGGCTGGTGGGCCTCAGTGGGGCTGCTGCCCTGAC
CCCCAGCACAATAAAATGAAACGTG

215/237

FIGURE 215

MRGSQEVLLMWLLVLAVGGTEHAYRPGRVCNAVRAHGDPVSESFVQRVYQPFLLTCDGHRAC
STYRTIYRTAYRRSPGLAPARPRYACCPGWKRTSGLPGACGAAICQPPCRNGGSCVQPGRCR
CPAGWRGDTCQSDVDECSARRGGCPQRCINTAGSYWCQCWEGHSLADGTLCKPKGGPPRVA
PNPTGVDSAMKEEVQLQSRVDLLEEKLQLVLAPLHSLASQALEHGLPDPGSLLVHSFQQLG
RIDSLSEQISFLEEQLGSCSCKKDS

Signal sequence:

1-19

216/237

FIGURE 216

CCACACGCGTCCGAAGCTGGCCCTGCACGCTGCAAGGGAGGCTCCTGTGGACAGGCCAGGCA
GGTGGGCCTCAGGAGGTGCCTCCAGGCAGGCCAGTGGCCTGAGGCCAGCAAGGGCTAGGG
TCCATCTCCAGTCCCAGGACACAGCAGCGGCCACCAGGCCACGCCTGGCTCCAGCAGCAT
CAGCAGCCCCCAGGACCGGGAGGCACAGGTGGCCCCACCACCCGGAGGAGCAGCTCCTGC
CCCTGTCCGGGGATGACTGATTCTCCTCCGCCAGGCCACCCAGAGGAGAAGGCCACCCGC
CTGGAGGCACAGGCCATGAGGGGCTCTCAGGAGGTGCTGCTGATGTGGCTTCTGGTGTGGC
AGTGGGCGGCACAGAGCACGCCTACCGGCCGGCGTAGGGTGTGCTGCTGTCCGGCTCACG
GGGACCTGTCTCCGAGTCGTTGTGCAGCGTGTGTAACAGCCCTCCTCACCAACCTGCGAC
GGGACCGGGCCTGCAGCACCTACCGAACCATCTATAGGACGCCCTACCGCCGCAGCCCTGG
GCTGGCCCCTGCCAGGCCCTGCTACCGTGCTGCCCGGCTGGAAGAGGACCAGCAGGGCTTC
CTGGGGCCTGTGGAGCAGCAATATGCCAGCCGCATGCCGAACGGAGGGAGCTGTGTCAG
CCTGGCCGCTGCCGCTGCCCTGCAGGATGGCGGGGTGACACTTGCAGTCAGATGTGGATGA
ATGCAGTGCTAGGAGGGGGCGGCTGTCCCAGCGCTGCGTCAACACCGCCGGCAGTTACTGGT
GCCAGTGTTGGGAGGGGCACAGCCTGTCTGCAGACGGTACACTCTGTGTGCCAAGGGAGGG
CCCCCCCAGGGTGGCCCCCAACCCGACAGGAGTGGACAGTGCAATGAAGGAAGAAGTGCAAGAG
GCTGCAGTCCAGGGTGGACCTGCTGGAGGAGAAGCTGCAGCTGGTGTGGCCCCACTGCACA
GCCTGGCCTCGCAGGCAGTGGAGCATGGGCTCCCAGGCCAGCCTCCTGGTGCAGTCC
TTCCAGCAGCTCGGCCGCATCGACTCCCTGAGCGAGCAGATTCTCCTGGAGGAGCAGCT
GGGGCCTGCTCCTGCAAGAAAGACTCGTGACTGCCAGGCCAGGCTGGACTGAGCCCC
TCACGCCGCCCTGCAGCCCCATGCCCTGCCAACATGCTGGGGTCCAGAACGCCACCTCG
GGGTGACTGAGCGGAAGGCCAGGCAGGGCCTTCCCTCTCCCTCCCTCGGGAG
GCTCCCCAGACCCCTGGCATGGGATGGGATGGGATCTTCTCTGTGAATCCACCCCTGGCTACC
CCCACCCCTGGCTACCCAAACGGCATCCAAAGGCCAGGTGGCCCTCAGCTGAGGGAAAGGTAC
GAGCTCCCTGCTGGAGGCTGGGACCCATGGCACAGGCCAGGCAGGCCAGGGCTGGGTGGGG
CCTCAGTGGGGCTGCTGCCCTGACCCCCAGCACAATAAAATGAAACGTG

217 / 237

FIGURE 217

MRGSQEVLLMWLLVLAGGTEHAYRPGRRVCAVRAHGDPVSESFVQRVYQPFLTTCDGHRAC
STYRTIYRTAYRRSPGLAPARPRYACCPGWKRTSGLPGACGAAICQPPCRNGGSCVQPGRCR
CPAGWRGDTQCSDVDECSARRGGCPQRCVNTAGSYWCQCWEGHSLAADGTLCPKGPPRVA
PNPTGVDSAMKEEVQRLQSRVDLLEEKLQLVLAPLHSLASQALEHGLPDPGSLLVHSFQQLG
RIDSLSEQISFLEEQLGSCSCKKDS

Signal sequence:

1-19

2181237

FIGURE 218

GGTTGCCACAGCTGGTTAGGGCCCCGACCACTGGGCCCCCTGTCAAGGAGGAGACAGCCTCCCGCCGGAG
GACAAGTCGCTGCCACCTTGGCTGCCGACGTGATTCCCTGGACGGTCCGTTCCCTGCCGTAGCTGCCGCCG
AGTTGGGTCTCCGTCTCAGGCCGGTCCCCCTCCTGGTCTCCCTCCTCCGCTGGCCGGTTATCGGAGG
AGATTGTCTCCAGGGCTAGCAATTGACTTTGATGATGTTGACCCAGCGGCAGGAATAGCAGGCAACGTGAT
TTCAAAGCTGGGCTCAGCCTCTGTTCTCTCGTGTAAATCGAAAACCATTGGAGCAGGAATTCCAATCA
TGTCGTGATGGTGGTGAGAAGAAGGTACACCGGAAACTGGAGAAACTCCCAGGAGAACACTTTGCTGTG
ATGGCCGCGTCAATGATGGCCGGAAAGGGCATTTTACCTGACCCCTTCTCATCCTGGGACATGTACAC
TCTTCTCGCCCTTGAGTGGCGTACCTGGCTGTCAGCTGCCATCCCTGATTGCTGCCATGCTCT
TCCCTTCTCCATGGCTACACTGTTGAGGACAGCTTCAGTGACCCCTGGAGTGTACCTCGGGCGCTACCAAGATG
AAGCAGCTTCATAGAAATGGAGATAGAAGCTACCAATGGTGGTGCCTGCCAGGGCCAGCGACCAACCGCTCGTA
TCAAGAATTTCAGATAAACAAACCAAGATTGTGAAACTGAAATACTGTTACACATGCAAGATCTCCGGCTCCCC
GGGCTCCCATGGCAGCATCTGTGACAACACTGTGTGGAGCGCTCGACCACACTGCCCTGGGTGGGAATTGTG
TTGAAAGAGGAAACTACCGCTACTTCTACCTCTCATCCTTCTCTCCCTCACAATCTATGTCTCGCCT
TCAACATCGTCTATGGCCCTCAAATCTTGAAATTGGCTTCTGGAGACATTGAAAGAAACTCTGGAACTG
TTCTAGAAGTCTCATTGCTTACACTCTGGCTGGGACTGACTGGATTTCATACCTTCTCGTGG
CTCTCAACAGACAACCAATGAAGACATCAAAGGATCATGGACAGGGAAAGATCGCGTCCAGAATCCCTACAGCC
ATGGCAATTGGTGAAGAACTGCTGTGAAGTGCTGTGGCCCTTGCCCCCAGTGTGCTGGATCGAAGGGGTA
TTTGCCACTGGAGGAAAGTGGAAAGTGCACCTCCAGTACTCAAGAGACAGTAGCAGCCTCTGCCACAGAGCC
CAGCCCCCACAGAACACCTGAACCAAATGAGATGCCGGAGGACAGCAGCACTCCGAAGAGATGCCACCTCCAG
AGCCCCCAGAGGCCACCACAGGAGGCAGCTGAAGCTGAGAAGTAGCCTATCTATGAAAGAGACTTTGTTGTGTT
TAATTAGGGCTATGAGAGATTTCAGGTGAGAAGTAAACCTGAGACAGAGCAAGTAAGCTGCCCTTTAACT
GTTTTCTTGGCTTTAGTCACCCAGGTGACACTGGCATTTCCTGCTGCAAGCTTTAAATTCTGAACT
CAAGGCAGTGCCAGAAGATGTCAGTCACCTCTGATAACTGGAAAATGGCTCTGGGCCCTGCCACTGGTCT
CCATGGCCTCAGCCACAGGGTCCCCCTGGACCCCTCTCTCCCTCCAGATCCCAGCCCTCTGCTGGGTCA
TGGTCTCATTCTGGGCTAAAAGTTTGAGACTGGCTCAAATCTCCAAAGCTGCTGCACGTGCTGAGTCCAGA
GGCAGTCACAGAGACCTGGCCAGGGATCTAAGTGGTTCTGGGTCTTCAGGACTGAAGAGGAGGGAGAG
TGGGGTCAGAAGATTCTCTGGCCACCAAGTGCAGCATTGCCACAAATCTTTAGGAATGGACAGGTACCT
TCCACTTGTGTAANNNNNNNNNNNNNNNNNNNNNNNNTGTTTCTTGTACTCTGCCCTTCTGACTCCCTCATTAGGAG
CAGGAATGGCAGTAATAAAAGTCTGCACTTGGTCAATTCTTCTCCCTCAGAGGAAGCCGAGTGTCACTTAAC
ACTATCCCCCTCAGACTCCCTGTGTGAGGCCCTGCAGAGGCCCTGAATGCACAAATGGAAACCAAGGCACAGAGAG
GCTCTCTCTCTCTCTCCCCGATGTACCCCTAAAAAAATGCTAACCAAGTCTTCCATTAAAGCT
CGGCTGAGTGAGGGAAAGCCAGCACTGCTGCCCTCTGGTAACTCACCTAAGGCCCTGCCACCTCTGGCT
ATGGTAACCACACTGGGGCTTCCCAAGCCCCCTCTCCAGCACTTCCACCGCAGAGTCCAGGCCACT
CACCTGGGGCTGGCTGGCCCCCAGTCAGCTGCTCAGGACCTGCTCTATTCAAGGAAGAAGATTATGT
ATTATATGTGGCTATATTCTAGAGCACCTGTGTTTCTCTTCAAGCCAGGGCTGTCTGGATGACTTAT
GCGGTGGGGAGTGTAAACCGGAACCTTCATCTATTGAAGGCAGTAAACTGTGCTAATGCA

2191237

FIGURE 219

MSVMVVRKKVTRKWEKLPGRNTFCCDGRVMMARQKGIFYLTFLFLILGTCTLFFAFECRYLAV
QLSPAIPVFAAMLFLFSMATLLRTSFSDPGVIPRALPDEAAIFIEMIEATNGAVPQGQRPPP
RIKNFQINNQIVKLKYCYTCKIFRPPRASHCSICDNCVERFDHHCPWVGNCVGKRNYRYFYL
FILSLSLLTIYVFANIVYVALKSLKIGFLETLKETPGTVLEVLICFFTLWSVVGLTGFHTF
LVALNQTTNEDIKGSGWTGKNRVQNPYSHGNIVKNCCEVLCGPLPPSVLDRRGILPLEESGR
PPSTQETSSSLPQSPAPTEHINSNEMPEDSSTPEEMPPPEPPEPQQEAAEAEK

Putative transmembrane domains:

amino acids 36-55 (type II TM), 65-84, 188-208, 229-245

220/237

FIGURE 220

AAAACCCGTATTTTACAATGCAAATAGACAATNANCCTGGAGGTCTTGAAATTAGGTAT
TATAGGGATGGTGGGTTGATTTNTCTGGAGGCTTGGCTTGGACTCTCNCTTCT
CCCACAGAGCNCTCGACCATCACTGCCCTGGTGGGAATTGTGGAAAGAGGAACTA
CCGCTANTCTACCTCTCATCCTTNTCTCTCCNCCTCACAAATCTATGTCTCGCCTCA
ACATCGT

227/237

FIGURE 221

GTTGTGTCCTCAGCAAAACAGTGGATTAAATCTCCTGCACAAGCTTGAGAGCAACACAA
TCTATCAGGAAAGAAAGAAAGAAAAACCGAACCTGACAAAAAAGAAGAAAAAGAAGAAGA
AAAAAAATCATGAAAACCATCCAGCAGCAAAATGCACAATTCTATCTCTGGGCAATCTCAC
GGGGCTGGCTGCTCTGTCTCTTCCAAGGAGTGCCGTGCGCAGCGGAGATGCCACCTTCC
CCAAAGCTATGGACAAACGTGACGGTCCGGCAGGGGGAGAGCGCCACCCTCAGGTGCACTATT
GACAACCGGGTACCCGGTGGCTGGCTAAACCGCAGCACCATCCTATGCTGGGAATGA
CAAGTGGTGCCTGGATCCTCGCGTGGTCTGAGCAACACCCAAACGCAGTACAGCATCG
AGATCCAGAACGTGGATGTATGACGAGGGCCCTTACACCTGCTCGGTGCAGACAGACAAAC
CACCAAAGACCTCTAGGGTCCACCTCATTGTGCAAGTATCTCCAAAATTGTAGAGATTTC
TTCAGATATCTCCATTAATGAAGGGAAACAATATTAGCCTCACCTGCATAGCAACTGGTAGAC
CAGAGCCTACGGTTACTTGGAGACACATCTCTCCAAAGCGGTTGGCTTGTGAGTGAAGAC
GAATACTTGGAAATTCAAGGCATCACCCGGAGCAGTCAGGGACTACGAGTGCAGTGCCTC
CAATGACGTGGCCGCGCCGTGGTACGGAGAGTAAAGGTACCGTGAACATCCACCATACA
TTTCAGAACCCAAGGGTACAGGTGTCCCCGTGGACAAAAGGGGACACTGCAGTGTGAAGCC
TCAGCAGTCCCCTCAGCAGAATTCCAGTGGTACAAGGATGACAAAAGACTGATTGAAGGAA
GAAAGGGGTGAAAGTGGAAACAGACCTTCCCTCTCAAAACTCATCTTCAATGTCTG
AACATGACTATGGAAACTACACTTGCCTGGCCCTCCAACAAGCTGGCCACACCAATGCCAGC
ATCATGCTATTGGTCCAGGCAGCGTCAAGCGAGGTGAGCAACGGCACGTGAGGGAGGGCAGG
CTCGTCTGGCTGCTGCCCTTCTGGCTTGCACCTGCTTCTCAAATTTGAGTGAGTGCC
ACTTCCCCACCCGGAAAGGCTGCCACCACCAACACAGCAATGGCAACAC
CGACAGCAACCAATCAGATATACAAATGAAATTAGAAGAAACACAGCCTCATGGGACAGA
AATTGAGGGAGGGGAACAAAGAATACTTGGGGGGAAAAGAGTTTAAAAAAGAAATTGAA
AATTGCCTTGCAGATATTAGGTACAATGGAGTTCTTCCAAACGGGAAGAACACAGC
ACACCCGGCTGGACCCACTGCAAGCTGCATCGTCAACCTCTTGGTGCAGTGTGGCAA
GGGCTCAGCCTCTGCCACAGAGTGCCACGTGGAACATTCTGGAGCTGCCATCCA
AATTCAATCAGTCCATAGAGACGAACAGAATGAGACCTCCGGCCAAGCGTGGCGTGC
GCACTTGGTAGACTGTGCCACCACGGCGTGTGAAACGTGAAATAAAAGAGCAAA
AAAAA

222/237

FIGURE 222

MKTIQPKMHNSISWAIFTGLAALCLFQGVPVRSGDATFPKAMDNVTVRQGESATLRCTIDNR
VTRVAWLNRSTILYAGNDKWCLDPRVVLLSNTQTQYSIEIQNVDVYDEGPYTCSVQTDNHPK
TSRVHLIVQVSPKIVEISSDISINEGNNISLTCAITGRPEPTVTWRHISPKAvgFVSEDEYL
EIQGITREQSGDYECASNDVAAPVVRRVKVTVNYPPIISEAKGTGPVGQKGTLQCEASAV
PSAEFQWYKDDKRLIEGKKGVKVENRPFLSKLIFFNVSEHDYGNYTCVASNLGHTNASIML
FGPGAVSEVNGTSRRAGCVWLLPLLVLHLLLKF

Signal peptide:

amino acids 1-28

223/237

FIGURE 223

GAAAAAAAATCATGAAAACCATCCAGCCAAAATGCACAATTCTATCTCTGGGCAATCTC
ACGGGGCTGGCTGCTCTGTCTCTCCAAGGAGTGCCC GTGCGCAGCGGAGATGCCACCTT
CCCCAAAGCTATGGACAACGTGACGGTCCGGCAGGGGGAGAGCGCCACCCTCAGGTGCACTA
TTGACAACCGGGTCACCCGGGTGGCTAAACCGCAGCACCATCCTCTATGCTGGGAAT
GACAAGTGGTGCCTGGATCCTCGCGTGGCCTTCTGAGCAACACCCAAACGCAGTACAGCAT
CGAGATCCAGAACGTGGATGTGTATGACGAGGGCCCTTACACCTGCTCGGTGCAGACAGACA
ACCACCCAAAGACCTCTAGGGTCCACCTCATTGTGCAAGTATCTCCAAAATTGTAGAGATT
TCTTCAGATATCTCCATTAATGAAGGAAACAATTAGCCTCACCTGCATAGCAAATGGTAG
ACCAGAG

224/237

FIGURE 224

ATGGCTGGTACGGCGGGCCGGCAGGGGACCGGGCGGGCCGGCCCAGCTGCCGGAGCCCTGA
ATCACCGCCTGGCCGACTCCACCATGAACGTGGCTGCAGGAGCTGGAGCTGGCAGCAACGTGGATTCCAG
AAGGGGACAAGACAGCTGTAGGCTCACGCACGCAGCTGGAGCTGGTCTAGCAGGTGCCTCTACTGCTGGCT
GCACTGCTCTGGGCTGCCTTGTGGCCTAGGGTCCAGTACACAGAGACCCATCCCACAGCACCTGCCTTACA
GAGGCTGCAATTGAGTGGCTGGAAAAATCCTGGAGTCCCTGGACCGAGGGTGAGGCCCTGTGAGGACTTTAC
CAGTCTCTGTGGGGCTGGATTGAGGAACCCCTGCCGATGGCGTTCTCGCTGGAACACCTTAAACAGC
CTCTGGGACAAAACAGGCCACTAAGGACCTGCTTGAAGGACCTGCTTGAACACCCACTTCAACTCCAGCAGTGAAGCTGAG
CAGAAGACACAGCCTACTTACCTTGCCTACAGGTGGAGCGCATTGAGGAGCTGGAGGCCAGCCACTGAGA
GACCTCATTGAGAAGATTGGTGGTGAACATTACGGGCCCTGGGACCAGGACAACATTGGAGGTGTTGAAG
GCAGTAGCAGGGACCTACAGGGCCACCCATTCTCACCGTACATCAGTGCCTAAGAGTTCCAACAGC
AATGTTATCCAGGTGGACAGTCTGGGCTCTTCGCCCTCTCGGGATTACTACTTAAACAGAACTGCCAATGAG
AAAGTGCTCACTGCCTATCTGGATTACATGGAGGAACACTGGGATGCTGCTGGGTGGCGGCCACCTCCACCGAGG
GAGCAGATGAGCAGGTGCTGGAGTGGAGATAACAGCTGGCAACATCACAGTGCCTGGAGGCCAGCAGCGCAG
GAGGAGAAGATCACCACAAAGATGAGCATTTCGGAGCTGCAGGCTCTGGCCCTCATGGACTGGCTTGAGTTC
CTGCTTCTTGCTGTCACCATGGAGTTGAGTGAATCTGAGCTGCACAAAGAGAAGCTGCTGGAGACCCCTATGGCACTAAGAAG
ACAAACCTCAAGGCTGGAGGAGCCTGGAGCTTGAAGTGCACAAAGAGAAGCTGCTGGAGACCCCTATGGCACTAAGAAG
TCCTGTGTGCCAGGGTGGCAGACCTGCATCTCAAACACGGATGACGCCCTGGCTTGCTTGGGTCACTCTC
GTGAAGGCCACGTTGACCGGCAAAGAAATTGAGAGGGATGATCAGCGAAATCCGGACCGCATTGAG
GAGGCCCTGGGACAGCTGGTTGGATGAGAAGACCCGCCAGGCAGGAAGGAGAAAGCAGATGCCATCTAT
GATGATTGGTTTCCAGACTTATCTGGAGGCCAAAGAGCTGGATGATGTTATGACGGGTACGAAATTCT
GAAGATTCTTCTTCAAACATGGTAATTGTAACACTCTCTGCAAGGTTATGGCTGACCAGCTCCGCAAG
CCTCCAGCCAGACAGTGGAGCATGACCCCCCAGACAGTGAATGCCTACACTTCAAACAGGAGATC
GTCTCCCCGCTGGCATCTGCAGGCCCTTCTATGCCGCAACCACCCAAAGGCCCTGAACCTGGTGGCATH
GGTGTGGTATGGCCATGAGTTGACGCATGCCATTGACCAAGGGCCGAGTATGACAAGAAGGGAACCTG
CGGCCCTGGTGGCAGAATGAGTCCCTGGCAGCCTCCGGAAACACAGGCCCTGCATGGAGGAACAGTACAATCAA
TACCAAGGTCAATGGGAGAGGCTCAACGGCCGCAAGCAGCTGGGGAGAACATTACTGACAACGGGGCTGAAG
GCTGCCCTACAATGCTTACAAGCATGGCTGAGAAAGCATGGGAGGAGCAGCAACTGCCAGCCGTGGCTCACC
AACCAACCAGCTTCTCTGGATTGGCTGGGACTGGTGGCTCGGTCCGGCACACCAGAGAGCTCAGGAGGG
CTGGTGACCGACCCCCACAGCCCTGGCCTGGCTCCCGACTCTCTCCAACTCCCGTGAATTCTGCC
CACTCGGCTGCCCTGTCGGCTCCCCCATGAACCCAGGGCAGCTGTGAGGTGTGGTAGACCTGGATCAGGGGA
GAAATGGCCAGCTGTCAACAGACCTGGGAGCTCTCTGACAAAGCTGTTGCTTGGGTTGGAGGAAGCAA
ATGCAAGCTGGCTGGGTCTAGTCCCTCCCCCCCACAGGTGACATGAGTACAGACCCCTCTCAATCACCACATTG
TGCCCTGCTTGGGGTGGCCCTGCCAGCAGAGCCCCACCATTCACTGTGACATCTTCCGTGTCACCC
GCCTGGAAGAGGTCTGGTGGGAGGCCAGTCCCCATAGGAAGGAGTCTGCC

225/237

FIGURE 225

MNVALQELGAGSNVGFKKGTRQLLGSRTQLELVLAGASLLAALLLGCLVALGVQYHRDPSH
STCLTEACIRVAGKILESLDRGVSPCEDFYQFSCGGWIRRNPPLPDGRSRWNTFNSLWDQNQA
ILKHLLENTTFNSSSEAEQKTQRFYLSCLQVERIEELGAQPLRDLIEKIGGWNIITGPWDQDN
FMEVLKAVAGTYRATPFFTYYISADSKSSNSNVIQVDQSGFLPLPSRDYYLNRTANEKVLTAY
LDYMEELGMILLGGRPTSTREQMQQVLELEIQLANITVPQDQRRDEEKIYHKMSISELQALAP
SMDWLEFLSFLLSPLLESDSEPVVVYGMDDYLQQVSELINRTEPSILNNYLIWNLVQKTTSSL
DRRFESAQEKLLETLYGKKSCVPRWQTCISNTDDALGFALGSLFVKATFDRQSKEIAEGMI
SEIRTAFFEEALGQLVWMDEKTRQAAKEKADAIYDMIGFPDFILEPKELDDVYDGYEISEDSF
FQNMLNLYNFSAKVMADQLRKPPSRDQWSMTPQTVNAYYLPTKNEIVFPAGILQAPFYARNH
PKALNFGGIGVVMGHELTHAFDDQGREYDKEGNLRPWWQNESLAAFRNHTACMEEQYNQYQV
NGERLNGRQTLGENITDNGGLKAAYNAYKAWLRKHGEEQQLPAVGLTNHQLFFVGFAQVWCS
VRTPESSHEGLVTDPHSPARFRVLGTLNSRDFLRHFGCPVGSPMNPGQLCEVV

Type II Transmembrane domain:

amino acids 32-57

226/237

FIGURE 226

GCCCGGGCCCTCCGCCCCCTCCGCACCTCCCGCCCTCCCTCCGCCGCCCCGTCCCGGCCCTCCTCCCTCCCTCC
 CAGCTGTCCCGTTCGCATGCCGAGCCTCCCGGCCGCCGCCAGAGCCCCCGTGTGCTCCCTCGGGCTGCTGCTGCT
 CGGCTCCCGGCCGCCGCCGCCGCCAGAGCCCCCGTGTGCTCCCTCGGGCTGAGAAGGAGGCCAGCGC
 CGTCGGGGAGCGGCCAGGTAGGTGGCGCCCGGGGAGGCGCGGGCGGGAGTCGGCTCGGGCTGGGGGGAGTCAGCGC
 CAGCCCGAGGGGGCGCGGGCGCAGGTGGCTCCGCAGGGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG
 GCGGTGCCTGGGACCCGGGACCCCGCAGGGCACACGGCGAGCTGGCAGCGGGCTCCAGGAGGGGGGGGGGGGGGGGG
 CAAGCCCGTCCCGCAGGTGTGACCTCTGGCGGAAGGTCTATGCCTTGAGCAGACGTGGCACCCGGACCTAGG
 GGAGCCATTGGGGTGTGCGCTGCTGTGCGAGGCGCAGTGGGGTGGCTGACCTAGGGGGGGGGGGGGGGGGGGGGGG
 CAGGGTCACTGCAAGAACATCAAAACAGATGCCCAACCCGGCTGTGGGAGCAGCGCCAGCTGCCGGGACA
 CTGCTGCCAGACACTGCCCGGAGACTCTGTGGCGCTGTGACAGGGCGAGCTGCCAGGGGGGGGGGGGGGGGGGGGG
 AGTCTCGCTGCTGCTCTAGCCTCCGCTTCTCTATCTCTACAGGCGGTGAGCAGGGGGGGGGGGGGGGGGGGGGGG
 CTCAGACTCCAATGGCAGTGTCTGTGAGCACCCTGCAAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG
 GCAGGGCAGTGCCTCGGGTCTCTGCCGCTCCTAGGGCAGAACAGAGCTGCACTGTTGACACTCACTCA
 CCCTCAGGGGGAGGTCTGGGGGCTCTCATCCGGCACCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG
 TCTAGAAGGG
 TTTTTGCTGCTCTTCGAGGGCTTCGAGGACTAACCCAGGGTCCAGGATCTACACCAGGGGGGGGGGGGGGGGG
 GCTACTGCGAGAACCTCAGGCCAATGTCTCAGGCCAAGGGCTTGCTGAGGTGCTGCCAACCTGACAGT
 CCAGGAGATGGACTGGCTGCTGCTGGGGGAGGCTGCAAGATGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG
 CAGTGGACACATTGTGCCAGGAAGAGCTGCCAGCTCTGCAAAGTGTCTTGTGGGGCTAATGCCCTGATCCC
 AGTCCAAACGGGTGCTGCCGGCTCAGGCCAGCCTCACTCTGCTAGGAAATGGCNCCCTGATCCTCAGGTGCAATT
 GGTAGGGACAACCAGTGAGGTGGGCTGACACTGGAAACCAAGCCTCAGCGGAGGGATCAGCCCACGTGCTCT
 GTGCCACATGGCTGGCTATCCTCCCTGCCCGGAGGGCTGGGTATCTGCCCTGGGCTGGGGGGGGGGGGGGGG
 TCATATGCTGCTGCAAGATGAGCTTCTCTGCAACGGACTTCCAGGGAGAGCTTCAGGGAGAGCTTGGGGGG
 ACAGTGGCTGCCCTGCCCTACTGTGGGGCATAGGCCCGCCCTGCCCTGAGCAGAGGCCCTGGTGTCTACC
 CCCTGTGAAGAGCCAAGCAGCAGGGCACGCCCTGGCTTCTGGATACCAACTGTCACTGCACTATGAAGTGT
 GCTGGCTGGGCTTGGTGGCTCAGAACAGGCACTGTCACTGCCACCTCTGGGCTCTGGAAACGCCAGGGGG
 TCGCGGCTGCTGAAGGGATTCTATGCTCAGAGGCCAGGGGTGTGGTGAAGGACCTGGAGGCCAGCTGCTGCG
 GCACCTGGCAAAGGCACTGGCTTCCCTGATGATCACCAACCAAGGTAAGGCCAGAGGGGGAGCTCCAGGGGG
 CTCTCTCCAGGTGCCATAGCCAACAAATGTGAGGTGGGGGACTGCCCTGGGAGGGGGGGGGGGGGGGGGGG
 GCGGGCGCTGGGGCTCCGGATAAGGCCCTGCTGCGGCCGCTGTGGTGTGGTCTCCGGGGGGGGGGGGGGGG
 CAAACCTGGTGGTCTGG
 TCGCTGGGCCAACACTAGGACCCGCTGCTCACTCTGCAACTGCCAGAGCACGAACGGTGTATCTGTGACCCGG
 GGTGTGCCACCAGGG
 TTTTGTGGTGAACGGGAGCTGGCGGGCAGCGGGTACCGGGTGGCACCCGGTGTGCCCCCTTGGCTTAATTAA
 GTGTGCTGCTGCAAGCAGGGGGGACTGGAGGGTGCAGTGTGAGAAGGTGCAAGTGTCCCCGGCTGGC
 CTGTGCCCAGCCTGTGCGTGTCAACCCCCACCAGACTGTCAGAACAGTGTCCAGGGTGGGGGGGGGGGGGG
 GGACCCCATGCAAGGCTGATGG
 CCCCTCAGTGCCCGGTTGGAGAGATGAGCTGATCACCTGCAAGATGGGGTAAGTGGGGAGCAGAGGGTTGT
 GTGAGGTGGGTACTGGGAGCCTGGTCTGGAGTAGGGAGACCTCCCAGGGAGGTCCCTGAAGAAGCTGAAGGTCA
 CTGTGCTCCAGTGCCTCTGGGGACACTCAGTGTCTGCTCTGTCTTGTACAGCAGGGGTGCCCTACTGTGAGC
 GGGATGACTGTTCACTGCCACTGTCTGTGGCTGGGGAGGAGACTGCACTGTGGAGTGGGGACTCTGTGATCAG
 GGCAGGGCTAAAGTGGGGAGGGAGGG
 GGG
 ACCTGGTGAATTGTTATTGACCTTTCTTACAATGAGATTCTGCAAGCTCAGAGAAATTAAAGCAACGGAG
 ATGAAGGTCAACCCAGCTGTGTGCACTGACCTGTTAGAAAATACTGGCCTTCTGGGACCAAGGCAGGGATGCT
 TGCCCTGCCCTATGCCCTCTGTGCCCTCTCCACTCCCTCTCCCTCCAAACATCCCTCCCTCTGTCTCC
 AGCAGCCCCAGAGACCAGACTGATCCAGAGCTGGAGAAAGAAGGCCAGGGCTTCTGGGAGCAGGCCAGGGGG
 AAGTGAACAGAGGGATGGGGCTGAGCTGGGGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG
 CCCAGTGGCTTGTCTCTGCTCTGCTCTGCTACTCCCTGCTACTCCCTGCTACTCCCTGCTACTCCCTGCT
 GAGAGGGAGCTGGGCCAGACCGAGGGTCAAGGCCACTCCAAGTCTGCCCTGCCACCCCTGGGCTCTGTCTGG
 GCCCCACCCCTTCTCTGTAACATAATGTCAGTGGCTTGGGGAGTTTAATTATCTTCACTCAGCACCAAG
 GGCCCCGGACACTCCACTCTGCTGCCCTGAGCTGAGCAGAGTCATTATGGAGAGTTGTATTAAAC
 ATTCTTTCACTTGGCATGAGGTTGGCTTTGTGGCCAGGAACCTGAGTGGGGCTGGTGGAGAAGGG
 GCNGAGAGTAGGGAGGTGAGAGAGAGGGAGCTCTGACACTTGGGGAGCTGAAAGAGACCTGGAGAGGGAG
 CGTGGCNNTTGGCTGGCATNCTGGGTCCCGCAGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG
 AATTAGGGAAAGTAGAACAGCAGGATTGACTCAAGTTAGTTAGCTTCCACATGCTGCCCTGTTGACTTCATG
 TTTGAAGTTGCTCCAGAGAGAGAATCAAGGTGTCACCGCCCCCTCTCCCTCCCTCCCTCCCTCCCT
 TTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT

227/237

FIGURE 227

GGCCGAGCGGGGGTGCCTGCGCGGCCGTGATGGCTGGTACGGCGGGCCGGCAGGGGA
CCGGGGCCGCCGGCCGGGAGCAGCTGCCGGAGCCCTGAATCACCGCCTGGCCCGAC
TCCACCATGAACGTCGCGCTGCAGGAGCTGGAGCTGGCAGCAACGTGGATTCCAGAAGGG
GACAAGACAGCTGTTAGGCTACGCACGCAGCTGGAGCTGGCTTAGCAGGTGCCTCTCTAC
TGCTGGCTGCACTGCTTCTGGCTGCCTTGTGGCCCTAGGGTCCAGTACCACAGAGACCCA
TCCCACAGCACCTGCCTTACAGAGGCCTGCATTGAGTGGCTGGAAAAATCCTGGAGTCCCT
GGACCGAGGGGTGAGCCCTGTGAGGACTTTACCAAGTTCTCCTGTGGGGCTGGATTCGGA
GGAACCCCCCTGCCGATGGCGTTCTCGCTGGAACACCTCAACAGCCTCTGGGACCAAAAC
CAGGCCATACTGAAGCACCTGCTGAAAACACCACCTTCAACTCCAGCAGTGAAGCTGAGCA
GAAGACACAGCGCTTACCTATCTGCCTACAGGTGGAGCGCATTGAGGAGCTGGGAGCCC
AGCCACTGAGAGACCTCATTGAGAAGATTGGTGGTGGAACATTACGGGGCCCTGGGACCAAG
GACAACTTATGGAGGTGTTGAAGGCAGTAGCAGGGACCTACAGGGCCACCCATTCTTCAC
CGTCTACATCAGTGCCACTCTAACAGTCAACAGCAATGTTATCCAGGTGGACCAGTCTG
GGCTTTCTGCCCTCTGGGATTACTACTTAAACAGAACTGCCAATGAGAAAGTAAGGAAC
ATCTTCCGAACCCCCATCCCTACCCCTGGCTGAGCTGGCTGATCCCTGTTGACTTTCCCT
TTGCCAAGGGTCAGAGCAGGGAAAGGTGAGCCTATCCTGTACCTAGTGAACAAACTGCCCT
CCTTCTTCTTCTTCTCCCTCCCTCCCTTCTCCCTTCTCCCTTCTCCCTTCCCT
TCTTATTCTCTAGTAGGTTCATAGACACCTACTGTGTGCCAGGTCCAGTGGGGAAATCG
GAGATATAAGTTCCGAGCCATTGCCACAGGAAGCGTTCAAGTGTGATGGGTTCATGGACCT
AGATAGGCTGATAACAAAGCTCACAAGAGGGCCTGAGGATTCAAGGAGAGACTTATGGAGCC
AGCAAAGTCTCCTGAAGAGATTGCATTGAGCCAGGTCTGTAG

2281237

FIGURE 228

ATGCCTACTACCTCCAACTAAGAACATGAGATCGTCTTCCCCGCTGGCATCCTGCAGGCC
TTCTATGCCCGCAACCACCCCAAGGCCCTGAACCTCGGTGGCATCGGTGTGGCATGGCCA
TGAGTTGACGCATGCCTTGATGACCAAGGGCGCGAGTATGACAAAGAACGGGAAACCTGCGGC
CCTGGTGGCAGAATGAGTCCCTGGCAGCCTCCGGAACCACACGGCCTGCATGGAGGAACAG
TACAATCAATACCAGGTCAATGGGAGAGGGCTCAACGGCCGCCAGACGCTGGGGAGAACAT
TGCTGACAACGGGGGGCTGAAGGCTGCCTACAATGCTTACAAAGCATGGCTGAGAAAGCATG
GGGAGGAGCAGCAACTGCCAGCCGTGGGCTCACCAACCACAGCTCTTCGTGGATT
GCCAGGTGTGGTGCTCGGTCCGCACACCAGAGAGCTCTCACGAGGGCTGGTGACCGACCC
CCACAGCCCTGCCCGCTTCCGCGTGCTGGCACTCTCTCCAACACTCCGTGACTTCCTGCGGC
ACTTCGGCTGCCCTGTCGGCTCCCCATGAACCCAGGGCAGCTGTGAGGTGTGGTAGACC
TGGATCAGGGGAGAAATGCCAGCTGTACCAAGACCTGGGCAGCTCTCCTGACAAAGCTGT
TTGCTCTGGGTGGGAGGAAGCAAATGCAAGCTGGCTGGGTCTAGTCCCTCCCCCCCACA
GGTACATGAGTACAGACCCCTCCTCAATCACACATTGTGCCTCTGCTTGGGGTGCCT
GCCTCCAGCAGAGCCCCCACCATTCACTGTGACATCTTCCGTGTCACCCCTGCCTGGAAGAG
GTCTGGGTGGGAGGCCAGTCCATAGGAAGGAGTCTGCCTCTGTCCCCAGGCTCACT
CAGCCTGGCGGCCATGGGCCTGCCGTGCCACTGTGACCCACAGGCCCTGGTG
TACCTCCTGGACTTCTCCCCAGGCTCACTCAGTGCACACTAGGGGTGGACTCAGCTCTGTC
TGGCTCACCCCTCACGGCTACCCCCACCTCACCCCTGTGCTCCTGTGCCACTGCTCCCAGTG
CTGCTGCTGACCTTCACTGACAGCTCCTAGTGGAAAGCCAAGGGCTCTGAAAGCCTCCTGC
TGCCCACGTTCCTGGCTGAGAGGGAAAGTCATATGTGATGCGGTACTGGTTCTGT
GTCTTAGGGACAAGCCTTAGCAAATGATTGATTCTCCCTGGACAAAGCAGGAAAGCAGATA
GAGCAGGGAAAAGGAAGAACAGAGTTATTTACAGAAAAGAGGGTGGAGGGTGTGGTCT
TGGCCCTTATAGGACC

229/237

FIGURE 229

CCACACGCGTCCGAGCCGCCGAGAATTAGACACACTCCGGACGCGGCCAAAAGCAACCGAGA
 GGAGGGGAGGCACAAACACCGAAAAACAAAAGAGAGAAACAACACCCAAACAACTGGGGTGG
 GGGGAAGAAAGAAAGAAAGAACCCACCCACCCACCAAAAAAAAAAAAAAAA
 AAAAAAAAATCCTGTGGCGCGCCTGGTCCCGGGAAAGACTCGCCAGCACCAGGGGG
 TGGGGGAGTGCAGCTGAAAGCTGCTGGAGAGTGAGCAGCCTAGCAGGGATGGAC**ATG**ATG
 CTGTTGGTGCAGGGTCTTGTGCTGAACCACTGGCTGGCGGGTGTCCCTCAGCCTGTG
 CTGCCTGCTACCCCTGCCTCCGGCTGGACAGAGTGTGGACTTCCCCTGGCGGCCGTGG
 ACAACATGATGGTCAGAAAAGGGACACGGCGGTGCTTAGGTGTTATTGGAAGATGGAGCT
 TCAAAGGGTGCCTGGTGAACCGGTCAAGTATTATTTGCGGGAGGTGATAAGTGGTCAGT
 GGATCCTCGAGTTCAATTCAACATTGAATAAAGGGACTACAGCCTCCAGATAACAGAATG
 TAGATGTGACAGATGATGGCCCATAACCGTGTCTGTTAGACTCAACATAACACCCAGAAC
 ATGCAGGTGCATCTAAGTGTGCAAGTTCCCTAAGATATGACATCTAAATGATATGAC
 CGTCAATGAAGGAACCAACGTCACTCTTACTTGTGACTGGAAACCAGAGCCTTCCA
 TTTCTGGCAGCACATCTCCCATCAGCAAAACCATTGAAAATGGACAATATTGGACATT
 TATGGAATTACAAGGGACCAGGCTGGGAATATGAATGCACTGGCTCTACTATTAGGAAATTAAAT
 CCCAGATGTGAGGAAAGTAAAGTTGTTGTCACATTGCTCTACTATTAGGAAATTAAAT
 CTGGCACCGTGACCCCCGGACCGAGTGGCCTGATAAGATGTGAAGGTGCAGGTGTGCCGCC
 CCAGCCTTGAATGGTACAAAGGAGAGAAGAAGCTTCAATGCCAACAAAGGAATTATTAT
 TCAAAATTAGCACAAGATCCATTCTCACTGTTACCAACGTGACACAGGAGCACTCGGCA
 ATTATACCTGTGTGGCTGCCAACAGCTAGGCACAACCAATGCGAGCCTGCCTTAACCT
 CCAAGTACAGCCCAGTATGGAATTACCGGGAGCGCTGATGTTCTTCTCTGCTGGTACCT
 TGTGTTGACACTGTCCTCTTCAACCAGCATATTCTACCTGAAGAATGCCATTCTACAAT**AAA**
 TTCAAAGACCCATAAAAGGCTTTAAGGATTCTGAAAGTGCTGATGGCTGGATCCAATCT
 GGTACAGTTGTTAAAGCAGCGTGGGATATAATCAGCAGTGCTTACATGGGATGATGCC
 TTCTGTAGAATTGCTCATTATGAAATACTTTAATTCTACTCTTTGATTAGCTACATTA
 CCTGTGAAGCAGTACACATTGTCCTTTTAAGACGTGAAAGCTGAAATTACTTTAG
 AGGATATTAATTGATTTCATGTTGAAATCTACAACCTTCAAAAGCATTAGTCATGGT
 CTGCTAGGTGCAAGGCTGAGTTACAAAACGAATATTGCACTGAATATGTGATTCTTAA
 GGCTGCAATACAAGCATTCAAGTCCCTGTTCAATAAGAGTCATCCACATTACAAGATG
 CATTTTTCTTTGATAAAAAGCAATAATATTGCTTCAGATTATTCTCAAATA
 TAACACATATCTAGATTTCTGCTGATGATATTCAAGGTTAGGAATGAGCCTGTAAT
 ATAACCTGGCTGTGCACTCTGCTCTCTTGTAAAGTTCAACTGAGCTGAAATGAGCT
 AATAATATTTCTCTTGTCTCAAACATAATATAAAATGTTGCTAAATCTACAATTG
 AAGTAAAATAAACCAAGAGTGAATCAAGTTAAACCATACACTATCTCTAAAGTAACGAAGGAGC
 TATTGGACTGTTAAACATCTTCTGCACTGACAATGGGTTGAGAATTGCCCCACACT
 AACCTAGTTCTGTGATGAGAGACAATTAAACAGTATAGTAAATATACCATATGATTC
 TTTAGTTGACTAAATGTTAGATCCACCGTGGAAATCATTCCCTTAAATGACAGCACA
 GTCCACTCAAAGGATTGCTAGCAATACAGCATCTTCTTCACTAGTCCAAGCCAAAAA
 TTTAAGATGATTGTCAGAAAGGGACAAAGTCTATCACCTAATATTACAAGAGTTGTA
 AGCGCTCATCTTAAATTGAGGCTGAGGTATTATGACAGTCGACCTGGAGGGTATGGA
 TATGGATATGGACGTTCCAGAGACTATAATGGCAGAAACCAGGGTGGTATGACCGCTACTC
 AGGAGGAATTACAGAGACAATTGACAACGTGAAATGAGACATGCACATAATATAGATACA
 CAAGGAATAATTCTGATCCAGGATCGTCCCTCCAAATGGCTGTATTATAAAGGTTTGG
 AGCTGCACTGAAGCATTAGACAGTTCCATCTTTAAATTGACCTGCCA
 AGTAGCTGAAGACCTTTAGACAGTTCCATCTTTAAATTGACCTGCCA
 AGACAAATTGGGACGTTGTCAAAAAAAAAAAAAAAA

230/237

FIGURE 230

MMLLVQGACCSNQWLAAVLLSLCCLLPSCLPAGQSVDFPWAADVNMVRKGDTAVLRCYLED
GASKGAWLNRSSIIIFAGGDKWSVDPRVSISTLNKRDSLQIQNVDTVDDGPYTCSVQTQHTP
RTMQVHLTVQVPPKIYDISNDMTVNEGTNVTLTCLATGKPEPSISWRHISPSAKPFENGQYL
DIYGITRDQAGEYECSAENAVSFDPDVRKVVVVFAPTIQEIKSGTVPGRSGLIRCEGAGV
PPPafeWYKGEKKLFNGQQGIIIQNFSTRSILTVTNVTQEHFGNYTCVAANKLGTNASLPL
NPPSTAQYGITGSADVLFSCWYLVLTSSFTSIFYLKNAILQ

Important features of the protein:**Signal peptide:**

amino acids 1-31

Transmembrane domain:

amino acids 326-345

N-glycosylation sites.

amino acids 71-75, 153-157, 273-277, 284-288, 292-296, 305-309

Casein kinase II phosphorylation site.

amino acids 147-151, 208-212, 224-228

Tyrosine kinase phosphorylation site.

amino acids 178-186

N-myristoylation sites.amino acids 7-13, 63-70, 67-73, 151-157, 239-245, 291-297,
302-308, 319-325**Myelin P0 protein:**

amino acids 92-121

231/237

FIGURE 231

AGTGGTCGATGGGAAGGATCTTCTCAAGTGGTCCTCTGAGGGGAGCATTCTGCTGG
CTCCAGGACTTGGCCATCTATAAGCTTGGCAATGGAGAATAAGAAAATTCTCAAGGAGGA
CGAGCTCTGAGTGAGACCCAACAAGCTGCTTTCACCAAATTGCAATGGAGCCTTCGAAA
TCAATGTTCAAAGCCAAAGAGGGAGAAATGGGTGAACCTCTCCCTAGCTGTGGTGGTCATC
TACCTGATCCTGCTACCGCTGGCGTGGCTGCTGGTGGTCAAGTCTGAATCTGCAGGC
GGGGCTCCGGGTCTGGAGATGTATTCCTCAATGACACTCTGGCGCTGAGGACAGCCCCT
CCTTCTCCTGCTGCAGTCAGCACACCCTGGAGAACACCTGGCTCAGGGTGCATCGAGGCTG
CAAGTCCTGCAGGCCAACTCACCTGGTCCCGTCAGCCATGAGCACTTGCTGCAGCGGGT
AGACAACCTCACTCAGAACCCAGGGATGTTAGAATCAAAGGTGAACAAGGCCCGGGCAGGTC
TTCAAGGTACAAGGGGGCATGGCATGCCTGGTGCCTGGCCCTGGCCCGGGACCACCTGCT
GAGAAGGGAGCCAAGGGGGCTATGGGACGAGATGGAGCAACAGGCCCTGGGACCCCCAAGG
CCCACCGGGAGTCAGGGAGAGGGGGCTCCAAGGACCCCAGGGTGCTCCAGGGAAAGCAAG
GAGCCACTGGCACCCCAGGACCCCAAGGAGAGAACGGCAGCAAAGGCATGGGGTCTCATT
GGCCCAAAAGGGAAACTGGAACTAAGGGAGAGAACGGAGACCTGGTCTCCAGGAAGCAA
AGGGGACAGGGGCATGAAAGGAGATGCAGGGTCAAGGGGCTCTGGGAGGCCAGGGAGTA
AAGGTGACTTCGGGAGGCCAGGCCACCCAGGTTGGCTGGTTCTGGAGCTAAAGGAGAT
CAAGGACAACCTGGACTGCAGGGTGTCCGGCCCTCCTGGTGCAGTGGACACCCAGGTGC
CAAGGGTGAGCCTGGCAGTGCTGGCTCCCTGGCAGCAGGACTTCCAGGGAGGCCGGGA
GTCCAGGAGGCCACAGGCCTGAAAGGAAGCAAAGGGGACACAGGACTCAAGGACAGCAAGGA
AGAAAAGGAGAACAGGAGTTCCAGGCCCTGCAGGTGTGAAGGGAGAACAGGGAGGCCAGG
GCTGGCAGGTCCCAAGGGAGGCCCTGGACAAGCTGGCAGAAGGGAGACCAGGGAGTGAAAG
GATCTTCTGGGAGCAAGGAGTAAAGGGAGAAAAGGTGAAAGAGGTGAAAACACTAGTGTCC
GTCAGGATTGTCGGCAGTAGTAACCGAGGCCGGCTGAAGTTACTACAGTGGTACCTGGG
GACAATTGCGATGACGAGTGGAAAATTCTGATGCCATTGTCTGCCGCATGCTGGGTT
ACTCCAAAGGAAGGCCCTGTACAAAGTGGAGCTGGCAGTGGCAGATCTGGCTGGATAAT
GTTCAGTGTGGGGCACGGAGAGTACCCCTGTGGAGCTGCACCAAGAACAGTGGGCCATCA
TGACTGCAGCCACGAGGAGGACGCAGGCAGGGTGCAGCGTCTGATGACCCGGAAACCCCTTCA
CTTCTCTGCTCCGAGGTGTCTGGCTCATATGTGGGAAGGCAGAGGATCTTGAGGAGT
TCCCTGGGACAACGTGAGCAGCCTCTGGAGAGGGGCCATTAATAAGCTCAACATCATTGA

232/237

FIGURE 232

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA68886
><subunit 1 of 1, 520 aa, 1 stop
><MW: 52658, pI: 9.16, NX(S/T): 3
MRNKKILKEDELLSETQQAAFHQIAMEPFEINVPKPKRRNGVNFSIAVVVIYLILLTAGAGL
LVVQVLNLQARLRVLEMYFLNDTLAAEDSPSFSLQSAHPGEHLAQGASRLQVLQAOLTWVR
VSHEHLLQRVDNFTQNPGMFRIKGEQGAPGLQGHKGAMGMPGAPGPPGPPAEKGAKGAMGRD
GATGPSGPQGPPGVKEAGLQGPQGAPGKQGATGTPGPQGEKGSKGDGGLIGPKGETGTKGE
KGDLGLPGSKGDGMKGDAGVMGPPGAQGSKGDFGRPGPPLAGFPGAKGDQQQPLQGVPG
PPGAVGHPGAKGEPGSAGSPGRAGLPGSPGSPGATGLKGSKGDTGLQGQQGRKGESGVPGPA
GVKGEQGSPLAGPKGAPGQAGQKGDQGVKGSSGEQGVKGEKGERGENSVRIVGSSNRGR
AEVYYSGTWGTICDDEWQNSDAIVFCRMLGYSKRALYKVGAGTGQIWLDNVQCRGTESTLW
SCTKNSWGHHDCSHEEDAGVECSV

Transmembrane domain:

amino acids 47-66 (type II)

N-glycosylation sites.

amino acids 43-47, 83-87, 136-140

Tyrosine kinase phosphorylation site.

amino acids 432-440

N-myristoylation sites.

amino acids 41-47, 178-184, 253-259, 274-280, 340-346, 346-352,
400-406, 441-447, 475-481, 490-496, 515-521

Amidation site.

amino acids 360-364

Leucine zipper pattern.

amino acids 56-78

Speract receptor repeat

amino acids 422-471, 488-519

C1q domain proteins.

amino acids 151-184, 301-334, 316-349

233/237

FIGURE 233

CCACACCGTCCGAAGGCAGACAAAGGTTCATTTGTAAGAAGCTCCTCCAGCACCTCCCT
CTTCTCCTTTGCCAAACTCACCCAGTGAGTGTGAGCATTAAAGAACATCCTCTGCCAAG
ACCAAAAGGAAAGAAGAAAAAGGGCAAAAGCCAAATGAAACTGATGGTACTGTTTCAC
CATTGGGCTAACCTTGCTGCTAGGAGTTCAAGCCATGCCTGCAAATCGCCTCTTGTACA
GAAAGATACTAAAGATCACAACTGTCACAACCTCCGGAAGGAGTAGCTGACCTGACACAG
ATTGATGTCAATGTCCAGGATCATTCTGGATGGGAAGGGATGTGAGATGATCTGTTACTG
CAAACCCAGCGAATTGCTCTGCTGCCAAAGACGTTTCTTGGACCAAGATCTCTTCG
TGATTCTTGCAACAATCAATGAGAATCTCATGTATTCTGGAGAACACCATTCTGATTTC
CCACAAACTGCACTACATCAGTATAACTGCATTCTAGTTCTATATAGTGCAATAGAGCAT
AGATTCTATAAATTCTTACTTGCTAAGACAAGTAAATCTGTGTTAAACAAGTAGTAATAAA
AGTTAATTCAATCTAAAAAAAAAAAAAA

234/237

FIGURE 234

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52758
<subunit 1 of 1, 98 aa, 1 stop
<MW: 11081, pI: 6.68, NX(S/T): 1
MKLMVLVFTIGLTLLGVQAMPANRLSCYRKILKDHNCNLPEGVADLTQIDVNQDHFWDG
KGCEMICYCNFSELLCCPKDVFFGPKISFVIPCNNQ

Important features:**Signal peptide:**

amino acids 1-20

N-glycosylation site.

amino acids 72-76

Tyrosine kinase phosphorylation site.

amino acids 63-71

235/237

FIGURE 235

CCACACGGTCCGGGACCGTGGCTGGACCCCAGGTCTGGAGCGAATTCCAGCCTGCAGGG
CTGATAAGCGAGGCATTAGTGAGATTGAGAGAGACTTACCCCGCCGTGGTGGAGGGC
GCGCAGTAGAGCAGCAGCACAGGCGCGGGTCCCAGGGAGGCCGGCTCTGCTCGCGCCGAG**ATG**
TGGAAATCTCCTCACGAAACCGACTCGGCTGTGGCCACCAGCGCCCGCGCTGGCTGTG
CGCTGGGGCGCTGGTGTGGCGGGTGGCTCTTCCTCTCGGCTTCCCTTCGGGTGGTTA
TAAAATCTCCAATGAAGCTACTAACATTACTCCAAAGCATAATATGAAAGCATTTGGAT
GAATTGAAAGCTGAGAACATCAAGAAGTCTTACATAATTTACACAGATAACCACATTAGC
AGGAACAGAACAAACTTCAGCTGCAAAGCAAATTCAATCCCAGTGGAAAGAATTGCC
TGGATTCTGAGCTAGCTATTGATGTCCTGTTCTACCCAAATAAGACTCATCCC
AACTACATCTCAATAATTAAATGAAGATGAAATGAGATTTCACACATCATTATTGAACC
ACCTCCTCCAGGATATGAAAATGTTCGGATATTGTACCACTTCAGTGCTTCTCTCCTC
AAGGAATGCCAGAGGGCAGTCACTGTATGTTAATGCACTGACGAACTGAAGACTTCTTAAA
TTGGAACGGGACATGAAAATCAATTGCTCTGGGAAAATTGTAATTGCCAGATATGGGAAAGT
TTTCAGAGGAATAAGGTTAAAATGCCAGCTGGCAGGGGCCAAAGGAGTCATTCTACT
CCGACCCCTGCTGACTACTTGCTCCTGGGTGAAGTCCTATCCAGACGGTTGGAATCTTCT
GGAGGTGGTGTCCAGCGTGGAAATATCTAAATCTGAATGGTGCAGGAGACCCCTCACACC
AGGTTACCCAGCAAATGAATATGTTAAGGCGTGGATTGCAAGAGGCTGTTGGTCTTCAA
GTATTCCCTGTTCATCCAATTGGATACTATGATGCAAGAGCTCCTAGAAAAAATGGGTGGC
TCAGCACCACAGATAGCAGCTGGAGAGGAAGTCCTAAAGTGCCTACAATGTTGGACCTGG
CTTACTGAAAATTTCTACACAAAAAGTCAGATGCACTCCACTTACCAATGAAGTGA
CGAGAATTACAATGTGATAGGTACTCTCAGAGGAGCAGTGGAACCAAGACAGATATGTCATT
CTGGGAGGTCACTGGGACTCATGGGTGTTGGTGTATTGACCTCTAGAGTGGAGCAGCTGT
TGTTCATGAAATTGTGAGGAGCTTGGAACACTGAAAAGGAAGGGTGGAGACCTAGAAGAA
CAATTGTTGCAAGCTGGGATGCAAGAGAATTGGTCTTGGTTACTGAGTGGCA
GAGGAGAATTCAAGACTCCTCAAGAGCGTGGCTTATATTAAATGCTGACTCATCTAT
AGAAGGAAACTACACTCTGAGAGTTGATGTACACCGCTGATGTACAGCTGGTACACAACC
TAACAAAAGAGCTGAAAAGCCCTGATGAAGGCTTGAAGGAAATCTTTATGAAAGTGG
ACTAAAAAAAGTCCTCCCCAGAGTCAGTGGCATGCCAGGATAAGCAAATTGGGATCTGG
AAATGATTGAGGTGTTCTCCAACGACTTGGATTGCTTCAAGGAGCAGGTATAACTA
AAAATTGGGAAACAAACAAATTCACTGAGGACTTCAGTGGCATGCCAGGATAAGCAAATTGGGATCTGG
GAGTTGGTGGAAAAGTTTATGATCCAATGTTAAATATCACCTCACTGTGGGCCAGGTG
AGGAGGGATGGTGTGAGCTAGCCAATTCCATAGTGTCTCCCTTGATGTGAGATTATG
CTGTAGTTTAAGAAAGTATGCTGACAAAATCTACAGTATTCTATGAAACATCCACAGGAA
ATGAAGACATACAGTGTATCTGATTGACTTCTGAGTAAAGAATTTCACAGAAAT
TGCTTCCAAGTTCAGTGAGAGACTCCAGGACTTGCACAAAAGCAACCAATAGTATTAAAGAA
TGATGAATGATCAACTCATGTTCTGAAAGAGCATTATTGATCCATTAGGGTTACCAAG
AGGCCTTTATAGGCATGTCATCTATGCTCCAAGCAGCCACAACAAGTATGCAGGGAGTC
ATTCCCAGGAATTATGATGCTCTGTTGATATTGAAAGCAAAGTGGACCCCTCCAAGGCCT
GGGAGAAGTGAAGAGACAGATTATGTTGCAGCCTCACAGTGCAGGAGCAGTCAGAGACT
TTGAGTGAAGTAGCCTAAAGGGATTAGAGAATCCGTATTGAATTGTGTTGATGTCA
CTCAGAAAGAATCGTAATGGGTATATTGATAAATTAAATTGGTATATTGAAATAAAGT
TGAATATTATATAA

236/237

FIGURE 236

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA52756
><subunit 1 of 1, 750 aa, 1 stop
><MW: 84305, pI: 6.93, NX(S/T): 10
MWNLHETDSAVATARRPRWLCAAGVLAGGFFLLGFLFGWFIKSSNEATNITPKHNMKAFL
DELKAENIKKFLHNFTQIPHLAGTEQNQQLAKQIQSQWKEFGLDSVELAHYDVLLSYPNKTH
PNYISIINEDGNEIFNTSLFEPPPGYENVSDIVPPFSAFSPQGMPEGDLVYVNYARTEDFF
KLERDMKINCSGKIVIARYGKVFRGNKVKNAQLAGAKGVILYSDPADYFAPGVKSYPDGWNL
PGGGVQRGNILNLNGAGDPLTPGYPANEYAYRRGIAEAVGLPSIPVHPIGYYDAQKLLEKMG
GSAPPDSSWRGSLKVPNVGPGFTGNFSTQKVHMHIHSTNEVTRIYNVIGTLRGAVEPDRYV
ILGGHRDSWVFGGIDPQSGAAVVHEIVRSFGTLKKEGWRPRRTILFASWDAEEFGLLGSTEW
AEENSRLLQERGVAYINADSSIENGTIRVDCTPLMYSLVHNLTKEKSPDEGFEGKSLYES
WTKKSPSPEFSGMPRISKLGSGNDFEVFFQRLGIASGRARYTKNWETNKFSGYPLYHSVYET
YELVEKFYDPMFKYHLTVAQVRGGMVFELANSIVLPFDCRDYAVVLRKYADKIYSISMKHPQ
EMKTYSVSFDSLFSAVKNTEIASKFSERLQDFDKSNPIVLRMMNDQLMFLERAFIDPLGLP
DRPFYRHVIYAPSSHNKYAGESFPGIYDALFDIESKVDPASKAWGEVKRQIYVAAFTVQAAA
TLSEVA

Signal sequence:

amino acids 1-40

N-glycosylation sites.

amino acids 76-80, 121-125, 140-144, 153-157, 195-199, 336-340,
459-463, 476-480, 638-642

Tyrosine kinase phosphorylation sites.

amino acids 363-372, 605-613, 606-613, 617-626

N-myristoylation sites.

amino acids 85-91, 168-174, 252-258, 256-262, 282-288, 335-341,
360-366, 427-433, 529-535, 707-713

237/237

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
14 September 2000 (14.09.2000)

PCT

(10) International Publication Number
WO 00/53756 A3

(51) International Patent Classification⁷: **C12N 15/12, 15/63, 5/10, 1/21, 1/19, C07K 14/47, 14/705, 19/00, 16/18, 16/28, G01N 33/68, 33/543, A61K 47/48, 51/08, 38/17**

(21) International Application Number: **PCT/US00/04341**

(22) International Filing Date: **18 February 2000 (18.02.2000)**

(25) Filing Language: **English**

(26) Publication Language: **English**

(30) Priority Data:

PCT/US99/05028	8 March 1999 (08.03.1999)	US
60/123,957	12 March 1999 (12.03.1999)	US
60/126,773	29 March 1999 (29.03.1999)	US
60/130,232	21 April 1999 (21.04.1999)	US
60/131,445	28 April 1999 (28.04.1999)	US
60/134,287	14 May 1999 (14.05.1999)	US
60/141,037	23 June 1999 (23.06.1999)	US
60/145,698	26 July 1999 (26.07.1999)	US
60/162,506	29 October 1999 (29.10.1999)	US
PCT/US99/28313	30 November 1999 (30.11.1999)	US
PCT/US99/28551	2 December 1999 (02.12.1999)	US
PCT/US99/28565	2 December 1999 (02.12.1999)	US
PCT/US99/30095	16 December 1999 (16.12.1999)	US
PCT/US99/31243	30 December 1999 (30.12.1999)	US
PCT/US99/31274	30 December 1999 (30.12.1999)	US
PCT/US00/00219	5 January 2000 (05.01.2000)	US
PCT/US00/00277	6 January 2000 (06.01.2000)	US
PCT/US00/00376	6 January 2000 (06.01.2000)	US

(71) Applicant (*for all designated States except US*): **GENENTECH, INC. [US/US]; 1 DNA Way, South San Francisco, CA 94080-4990 (US).**

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): **ASHKENAZI, Avi, J. [US/US]; 1456 Tarrytown Street, San Mateo, CA 94402 (US). BAKER, Kevin, P. [GB/US]; 14006 Indian Run Drive, Darnestown, MD 20878 (US). BOTSTEIN, David**

[US/US]; 2539 Somerset Drive, Belmont, CA 94002 (US). DESNOYERS, Luc [CA/US]; 2050 Stockton Street, San Francisco, CA 94133 (US). EATON, Dan, L. [US/US]; 75 Knight Drive, San Rafael, CA 94901 (US). FERRARA, Napoleone [US/US]; 2090 Pacific Avenue #704, San Francisco, CA 94109 (US). FILVAROFF, Ellen [US/US]; 538 18th Avenue, San Francisco, CA 94121 (US). FONG, Sherman [US/US]; 19 Basinside Way, Alameda, CA 94502 (US). GAO, Wei-Qiang [CN/US]; 641 Pilgrim Drive, Foster City, CA 94404 (US). GERBER, Hanspeter [CH/US]; 1121 Tennessee Street #5, San Francisco, CA 94107 (US). GERRITSEN, Mary, E. [CA/US]; 541 Parrott Drive, San Mateo, CA 94402 (US). GODDARD, Audrey [CA/US]; 110 Congo Street, San Francisco, CA 94131 (US). GODOWSKI, Paul, J. [US/US]; 2627 Easton Drive, Burlingame, CA 94010 (US). GRIMALDI, Christopher, J. [US/US]; 1434 36th Avenue, San Francisco, CA 94122 (US). GURNEY, Austin, L. [US/US]; 1 Debbie Lane, Belmont, CA 94002 (US). HILLAN, Kenneth, J. [GB/US]; 64 Seward Street, San Francisco, CA 94114 (US). KLJAVIN, Ivar, J. [US/US]; 3673 Crescent Drive, Lafayette, CA 94549 (US). KUO, Sophia, S. [US/US]; Apartment 3, 59 Surrey Street, San Francisco, CA 94131 (US). NAPIER, Mary, A. [US/US]; 1015 Hayne Road, Hillsborough, CA 94010 (US). PAN, James [CA/US]; 2705 Coronet Boulevard, Belmont, CA 94002 (US). PAONI, Nicholas, F. [US/US]; 1756 Terrace Drive, Belmont, CA 94002 (US). ROY, Margaret, Ann [US/US]; 2960 Webster Street #4, San Francisco, CA 94123 (US). SHELTON, David, L. [US/US]; 5845 Clover Drive, Oakland, CA 94618 (US). STEWART, Timothy, A. [US/US]; 465 Douglass Street, San Francisco, CA 94114 (US). TUMAS, Daniel [US/US]; 3 Rae Avenue, Orinda, CA 94563 (US). WILLIAMS, P., Mickey [US/US]; 509 Alto Avenue, Half Moon Bay, CA 94019 (US). WOOD, William, I. [US/US]; 35 Southdown Court, Hillsborough, CA 94010 (US).

(74) Agents: **KRESNAK, Mark, T. et al.; Genentech, Inc., 1 DNA Way, South San Francisco, CA 94080-4990 (US).**

(81) Designated States (*national*): **AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.**

[Continued on next page]

(54) Title: **SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC ACIDS ENCODING THE SAME**

(57) Abstract: The present invention is directed to novel polypeptides and to nucleic acid molecules encoding those polypeptides. Also provided herein are vectors and host cells comprising those nucleic acid sequences, chimeric polypeptide molecules comprising the polypeptides of the present invention fused to heterologous polypeptide sequences, antibodies which bind to the polypeptides of the present invention and to methods for producing the polypeptides of the present invention.

WO 00/53756 A3



(84) **Designated States (regional):** ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

(88) **Date of publication of the international search report:**
1 February 2001

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

Published:

— *With international search report.*

INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 00/04341

A. CLASSIFICATION OF SUBJECT MATTER				
IPC 7	C12N15/12	C12N15/63	C12N5/10	C12N1/21
	C07K14/47	C07K14/705	C07K19/00	C07K16/18
	G01N33/68	G01N33/543	A61K47/48	A61K51/08

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 C12N C07K G01N A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 98 57983 A (ZYMOGENETICS INC) 23 December 1998 (1998-12-23) the whole document ---	1-3, 5-13, 15-21
A	KLEIN R D ET AL: "Selection for genes encoding secreted proteins and receptors" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, US, NATIONAL ACADEMY OF SCIENCE. WASHINGTON, no. 93, 1 July 1996 (1996-07-01), pages 7108-7113, XP002077277 ISSN: 0027-8424 the whole document ---	-/-

Further documents are listed in the continuation of box C.

Patent family members are listed in annex.

* Special categories of cited documents :

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search

Date of mailing of the international search report

20 July 2000

18.10.00

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Smalt, R

INTERNATIONAL SEARCH REPORT

International Application No PCT/US 00/04341

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	<p>YOKOYAMA-KOBAYASHI M ET AL: "A signal sequence detection system using secreted protease activity as an indicator" GENE,NL,ELSEVIER BIOMEDICAL PRESS. AMSTERDAM, vol. 163, no. 2, 3 October 1995 (1995-10-03), pages 193-196, XP004041983 ISSN: 0378-1119 the whole document ---</p>	
A	<p>EP 0 834 563 A (SMITHKLINE BEECHAM CORP) 8 April 1998 (1998-04-08) the whole document ---</p>	
A	<p>WO 97 07198 A (GENETICS INST) 27 February 1997 (1997-02-27) the whole document ---</p>	
P,X, L	<p>WO 99 46281 A (BAKER KEVIN P ;CHEN JIAN (US); GENENTECH INC (US); GURNEY AUSTIN () 16 September 1999 (1999-09-16) whole document, particularly the pasasages referring to PRO213 (e.g. pages 2,50,123,183), and the claims ---</p>	1-3, 5-13, 15-21
P,X	<p>WO 99 54437 A (MILLENNIUM BIOTHERAPEUTICS INC) 28 October 1999 (1999-10-28) the whole document -----</p>	1-3, 5-13, 15-21

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 00/04341

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. Claims Nos.: because they relate to subject matter not required to be searched by this Authority, namely:

2. Claims Nos.: because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

3. Claims Nos.: because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Invention 1 : claims 1-3, 5-13, 15-21, all partially

Remark on Protest

- The additional search fees were accompanied by the applicant's protest.
 No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Invention 1: claims 1-3,5-13,15-21, all partially

Nucleic acid with seq.ID.1, encoding a polypeptide comprising the amino acid sequence as represented in seq.ID.2 or a nucleic acid having at least 80% homology thereto, vector comprising said nucleic acid, host cell comprising said vector, process for producing the protein of seq.ID.2 using said host, the isolated protein or one having at least 80% homology thereto, a chimeric protein of said peptide fused to a heterologous sequence, isolated extracellular domain of said protein or said protein lacking its signal peptide, and an antibody against said polypeptide.

Invention 2: claims 1-29,38-43, and 50-53,
all partially

Nucleic acid with seq.ID.611, encoding a polypeptide comprising the amino acid sequence as represented in seq.ID.612 or a nucleic acid having at least 80% homology thereto, vector comprising said nucleic acid, host cell comprising said vector, process for producing the protein of seq.ID.612 using said host, the isolated protein or one having at least 80% homology thereto, a chimeric protein of said peptide fused to a heterologous sequence, isolated extracellular domain of said protein or said protein lacking its signal peptide, and an antibody against said polypeptide. Also a method of detecting PRO337 (seq.ID.523) using its interaction with PRO4993 (seq.ID.612), method for linking a bioactive molecule to a cell expressing PRO337 through the use of PRO4993, and method of modulating at least one activity of said cell thereby.

Invention 3: claims 1-29,38-43, and 50-53,
all partially

Nucleic acid with seq.ID.522, encoding a polypeptide comprising the amino acid sequence as represented in seq.ID.523 or a nucleic acid having at least 80% homology thereto, vector comprising said nucleic acid, host cell comprising said vector, process for producing the protein of seq.ID.523 using said host, the isolated protein or one having at least 80% homology thereto, a chimeric protein of said peptide fused to a heterologous sequence, isolated extracellular domain of said protein or said protein lacking its signal peptide, and an antibody against said polypeptide. Also a method of detecting PRO4993 (seq.ID.612) using its interaction with PRO337 (seq.ID.523), method for linking a bioactive molecule to a cell expressing PRO4993 through the use of PRO337, and method of modulating at least one activity of said cell thereby.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Invention 4: claims 1-21,30-37,44-49, and 54-57,
all partially

Nucleic acid with seq.ID.613, encoding a polypeptide comprising the amino acid sequence as represented in seq.ID.614 or a nucleic acid having at least 80% homology thereto, vector comprising said nucleic acid, host cell comprising said vector, process for producing the protein of seq.ID.614 using said host, the isolated protein or one having at least 80% homology thereto, a chimeric protein of said peptide fused to a heterologous sequence, isolated extracellular domain of said protein or said protein lacking its signal peptide, and an antibody against said polypeptide. Also a method of detecting PRO700 (seq.ID.523), PRO725 (seq.ID.616) and/or PRO739 (seq.ID.618) using their interaction with PRO1559 (seq.ID.614), method for linking a bioactive molecule to a cell expressing PRO700, PRO725 or PRO739 through the use of PRO1559, and method of modulating at least one activity of said cell thereby.

Invention 5: claims 1-21,30-37,44-49, and 54-57,
all partially

Nucleic acid with seq.ID.89, encoding a polypeptide comprising the amino acid sequence as represented in seq.ID.90 or a nucleic acid having at least 80% homology thereto, vector comprising said nucleic acid, host cell comprising said vector, process for producing the protein of seq.ID.90 using said host, the isolated protein or one having at least 80% homology thereto, a chimeric protein of said peptide fused to a heterologous sequence, isolated extracellular domain of said protein or said protein lacking its signal peptide, and an antibody against said polypeptide. Also a method of detecting PRO1559 (seq.ID.614) using its interaction with PRO700 (seq.ID.90), method for linking a bioactive molecule to a cell expressing PRO1559 through the use of PRO700, and method of modulating at least one activity of said cell thereby.

Invention 6: claims 1-21,30-37,44-49, and 54-57,
all partially

Nucleic acid with seq.ID.615, encoding a polypeptide comprising the amino acid sequence as represented in seq.ID.616 or a nucleic acid having at least 80% homology thereto, vector comprising said nucleic acid, host cell comprising said vector, process for producing the protein of seq.ID.616 using said host, the isolated protein or one having at least 80% homology thereto, a chimeric protein of

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

PRO381, as represented by seq.ID's 144 and 145,
PRO386, as represented by seq.ID's 149 and 150,
PRO540, as represented by seq.ID's 156 and 157,
PRO615, as represented by seq.ID's 161 and 162,
PRO618, as represented by seq.ID's 168 and 169,
PRO719, as represented by seq.ID's 177 and 178,
PRO724, as represented by seq.ID's 182 and 183,
PRO772, as represented by seq.ID's 189 and 190,
PRO852, as represented by seq.ID's 195 and 196,
PRO853, as represented by seq.ID's 205 and 206,
PRO860, as represented by seq.ID's 210 and 211,
PRO846, as represented by seq.ID's 215 and 216,
PRO862, as represented by seq.ID's 220 and 221,
PRO864, as represented by seq.ID's 225 and 226,
PRO792, as represented by seq.ID's 230 and 231,
PRO866, as represented by seq.ID's 235 and 236,
PRO871, as represented by seq.ID's 244 and 245,
PRO873, as represented by seq.ID's 253 and 254,
PRO940, as represented by seq.ID's 258 and 259,
PRO941, as represented by seq.ID's 263 and 264,
PRO944, as represented by seq.ID's 269 and 270,
PRO983, as represented by seq.ID's 283 and 284,

Claim : 9

PRO1057, as represented by seq.ID's 295 and 296,
PRO1071, as represented by seq.ID's 300 and 301,
PRO1072, as represented by seq.ID's 302 and 303,
PRO1075, as represented by seq.ID's 308 and 309,
PRO181, as represented by seq.ID's 321 and 322,
PRO195, as represented by seq.ID's 329 and 330,
PRO865, as represented by seq.ID's 336 and 337,
PRO827, as represented by seq.ID's 345 and 346,
PRO1114, as represented by seq.ID's 351 and 352,
PRO237, as represented by seq.ID's 357 and 358,
PRO541, as represented by seq.ID's 362 and 363,
PRO273, as represented by seq.ID's 369 and 370,
PRO701, as represented by seq.ID's 374 and 375,
PRO704, as represented by seq.ID's 379 and 380,
PRO706, as represented by seq.ID's 384 and 385,
PRO707, as represented by seq.ID's 389 and 390,
PRO322, as represented by seq.ID's 394 and 395,
PRO526, as represented by seq.ID's 399 and 400,
PRO531, as represented by seq.ID's 404 and 405,
PRO534, as represented by seq.ID's 409 and 410,
PRO697, as represented by seq.ID's 414 and 415,
PRO717, as represented by seq.ID's 419 and 420,
PRO731, as represented by seq.ID's 424 and 425,
PRO218, as represented by seq.ID's 429 and 430,
PRO768, as represented by seq.ID's 436 and 437,
PRO771, as represented by seq.ID's 441 and 442,
PRO733, as represented by seq.ID's 446 and 447,
PRO162, as represented by seq.ID's 451 and 452,
PRO788, as represented by seq.ID's 453 and 454,

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

PRO1008, as represented by seq.ID's 455 and 456,
PRO1012, as represented by seq.ID's 458 and 459,
PRO1014, as represented by seq.ID's 463 and 464,
PRO1017, as represented by seq.ID's 465 and 466,
PRO474, as represented by seq.ID's 467 and 468,
PRO1031, as represented by seq.ID's 469 and 470,
PRO938, as represented by seq.ID's 471 and 472,
PRO1082, as represented by seq.ID's 476 and 477,
PRO1083, as represented by seq.ID's 482 and 483,
PRO200, as represented by seq.ID's 487 and 488,
PRO285, as represented by seq.ID's 495 and 496,
PRO286, as represented by seq.ID's 497 and 498
PRO213-1, as represented by seq.ID's 505 and 506,
PRO1330, as represented by seq.ID's 507 and 508,
PRO1449, as represented by seq.ID's 509 and 510,
PRO298, as represented by seq.ID's 514 and 515,
PRO403, as represented by seq.ID's 525 and 526,

For the sake of conciseness, the first subject matter is explicitly defined, the other subject matters are defined by analogy thereto.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/US 00/04341

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9857983	A 23-12-1998	AU 7979898 A EP 0996628 A	04-01-1999 03-05-2000
EP 0834563	A 08-04-1998	JP 10179178 A US 5824504 A	07-07-1998 20-10-1998
WO 9707198	A 27-02-1997	US 5707829 A AU 6712396 A AU 6768596 A CA 2227220 A CA 2229208 A EP 0839196 A EP 0851875 A JP 11510045 T US 6043344 A WO 9704097 A US 6074849 A US 5969093 A	13-01-1998 18-02-1997 12-03-1997 06-02-1997 27-02-1997 06-05-1998 08-07-1998 07-09-1999 28-03-2000 06-02-1997 13-06-2000 19-10-1999
WO 9946281	A 16-09-1999	AU 3072199 A AU 3075099 A WO 9947677 A AU 1532499 A EP 1032667 A WO 9927098 A AU 3757099 A WO 9954467 A AU 1070399 A EP 1025227 A WO 9920756 A	27-09-1999 11-10-1999 23-09-1999 15-06-1999 06-09-2000 03-06-1999 08-11-1999 28-10-1999 10-05-1999 09-08-2000 29-04-1999
WO 9954437	A 28-10-1999	AU 3965799 A	08-11-1999